EXHIBIT A-1

UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

NOTICE OF ALLOWANCE AND FEE(S) DUE

108676 7590 12/22/2021	
Ronald M. Kachmarik	
Cooper Legal Group LLC	
1388 Ridge Road, Unit 1	
Hinckley, OH 44233	

EXAMINER
AZARIAN, SEYED H

ART UNIT	PAPER NUMBER
2667	

DATE MAILED: 12/22/2021

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
16/031,125	07/10/2018	Philippe SALAH	N&P-51423	3171

TITLE OF INVENTION: METHOD FOR ANALYZING AN IMAGE OF A DENTAL ARCH

APPLN. TYPE	ENTITY STATUS	ISSUE FEE DUE	PUBLICATION FEE DUE	PREV. PAID ISSUE FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	UNDISCOUNTED	$1200	$0.00	$0.00	$1200	03/22/2022

THE APPLICATION IDENTIFIED ABOVE HAS BEEN EXAMINED AND IS ALLOWED FOR ISSUANCE AS A PATENT. PROSECUTION ON THE MERITS IS CLOSED. THIS NOTICE OF ALLOWANCE IS NOT A GRANT OF PATENT RIGHTS. THIS APPLICATION IS SUBJECT TO WITHDRAWAL FROM ISSUE AT THE INITIATIVE OF THE OFFICE OR UPON PETITION BY THE APPLICANT. SEE 37 CFR 1.313 AND MPEP 1308.

THE ISSUE FEE AND PUBLICATION FEE (IF REQUIRED) MUST BE PAID WITHIN THREE MONTHS FROM THE MAILING DATE OF THIS NOTICE OR THIS APPLICATION SHALL BE REGARDED AS ABANDONED. THIS STATUTORY PERIOD CANNOT BE EXTENDED. SEE 35 U.S.C. 151. THE ISSUE FEE DUE INDICATED ABOVE DOES NOT REFLECT A CREDIT FOR ANY PREVIOUSLY PAID ISSUE FEE IN THIS APPLICATION. IF AN ISSUE FEE HAS PREVIOUSLY BEEN PAID IN THIS APPLICATION (AS SHOWN ABOVE), THE RETURN OF PART B OF THIS FORM WILL BE CONSIDERED A REQUEST TO REAPPLY THE PREVIOUSLY PAID ISSUE FEE TOWARD THE ISSUE FEE NOW DUE.

HOW TO REPLY TO THIS NOTICE:

I. Review the ENTITY STATUS shown above. If the ENTITY STATUS is shown as SMALL or MICRO, verify whether entitlement to that entity status still applies.

If the ENTITY STATUS is the same as shown above, pay the TOTAL FEE(S) DUE shown above.

If the ENTITY STATUS is changed from that shown above, on PART B - FEE(S) TRANSMITTAL, complete section number 5 titled "Change in Entity Status (from status indicated above)".

For purposes of this notice, small entity fees are 1/2 the amount of undiscounted fees, and micro entity fees are 1/2 the amount of small entity fees.

II. PART B - FEE(S) TRANSMITTAL, or its equivalent, must be completed and returned to the United States Patent and Trademark Office (USPTO) with your ISSUE FEE and PUBLICATION FEE (if required). If you are charging the fee(s) to your deposit account, section "4b" of Part B - Fee(s) Transmittal should be completed. If an equivalent of Part B is filed, a request to reapply a previously paid issue fee must be clearly made, and delays in processing may occur due to the difficulty in recognizing the paper as an equivalent of Part B.

III. All communications regarding this application must give the application number. Please direct all communications prior to issuance to Mail Stop ISSUE FEE unless advised to the contrary.

IMPORTANT REMINDER: Maintenance fees are due in utility patents issuing on applications filed on or after Dec. 12, 1980. It is patentee's responsibility to ensure timely payment of maintenance fees when due. More information is available at www.uspto.gov/PatentMaintenanceFees.

Page 1 of 3

PTOL-85 (Rev. 02/11)

PART B - FEE(S) TRANSMITTAL

Complete and send this form, together with applicable fee(s), by mail or fax, or via EFS-Web.

By mail, send to: Mail Stop ISSUE FEE
Commissioner for Patents
P.O. Box 1450
Alexandria, Virginia 22313-1450

By fax, send to: (571)-273-2885

INSTRUCTIONS: This form should be used for transmitting the ISSUE FEE and PUBLICATION FEE (if required). Blocks 1 through 5 should be completed where appropriate. All further correspondence including the Patent, advance orders and notification of maintenance fees will be mailed to the current correspondence address as indicated unless corrected below or directed otherwise in Block 1, by (a) specifying a new correspondence address; and/or (b) indicating a separate "FEE ADDRESS" for maintenance fee notifications.

CURRENT CORRESPONDENCE ADDRESS (Note: Use Block 1 for any change of address)

108676   7590   12/22/2021

Ronald M. Kachmarik
Cooper Legal Group LLC
1388 Ridge Road, Unit 1
Hinckley, OH 44233

Note: A certificate of mailing can only be used for domestic mailings of the Fee(s) Transmittal. This certificate cannot be used for any other accompanying papers. Each additional paper, such as an assignment or formal drawing, must have its own certificate of mailing or transmission.

Certificate of Mailing or Transmission
I hereby certify that this Fee(s) Transmittal is being deposited with the United States Postal Service with sufficient postage for first class mail in an envelope addressed to the Mail Stop ISSUE FEE address above, or being transmitted to the USPTO via EFS-Web or by facsimile to (571) 273-2885, on the date below.

_____ (Typed or printed name)

_____ (Signature)

_____ (Date)

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
16/031,125	07/10/2018	Philippe SALAH	N&P-51423	3171

TITLE OF INVENTION: METHOD FOR ANALYZING AN IMAGE OF A DENTAL ARCH

APPLN. TYPE	ENTITY STATUS	ISSUE FEE DUE	PUBLICATION FEE DUE	PREV. PAID ISSUE FEE	TOTAL FEE(S) DUE	DATE DUE
nonprovisional	UNDISCOUNTED	$1200	$0.00	$0.00	$1200	03/22/2022

EXAMINER	ART UNIT	CLASS-SUBCLASS
AZARIAN, SEYED H	2667	382-128000

1. Change of correspondence address or indication of "Fee Address" (37 CFR 1.363).

❏ Change of correspondence address (or Change of Correspondence Address form PTO/AIA/122 or PTO/SB/122) attached.

❏ "Fee Address" indication (or "Fee Address" Indication form PTO/AIA/47 or PTO/SB/47; Rev 03-02 or more recent) attached. **Use of a Customer Number is required.**

2. For printing on the patent front page, list
(1) The names of up to 3 registered patent attorneys or agents OR, alternatively,
(2) The name of a single firm (having as a member a registered attorney or agent) and the names of up to 2 registered patent attorneys or agents. If no name is listed, no name will be printed.

1 _____

2 _____

3 _____

3. ASSIGNEE NAME AND RESIDENCE DATA TO BE PRINTED ON THE PATENT (print or type)

PLEASE NOTE: Unless an assignee is identified below, no assignee data will appear on the patent. If an assignee is identified below, the document must have been previously recorded, or filed for recordation, as set forth in 37 CFR 3.11 and 37 CFR 3.81(a). Completion of this form is NOT a substitute for filing an assignment.

(A) NAME OF ASSIGNEE   (B) RESIDENCE: (CITY and STATE OR COUNTRY)

Please check the appropriate assignee category or categories (will not be printed on the patent) : ❏ Individual ❏ Corporation or other private group entity ❏ Government

4a. Fees submitted: ❏ Issue Fee   ❏ Publication Fee (if required)   ❏ Advance Order - # of Copies _____

4b. Method of Payment: *(Please first reapply any previously paid fee shown above)*

❏ Electronic Payment via EFS-Web   ❏ Enclosed check   ❏ Non-electronic payment by credit card (Attach form PTO-2038)

❏ The Director is hereby authorized to charge the required fee(s), any deficiency, or credit any overpayment to Deposit Account No. _____

5. **Change in Entity Status** (from status indicated above)

❏ Applicant certifying micro entity status. See 37 CFR 1.29

❏ Applicant asserting small entity status. See 37 CFR 1.27

❏ Applicant changing to regular undiscounted fee status.

NOTE: Absent a valid certification of Micro Entity Status (see forms PTO/SB/15A and 15B), issue fee payment in the micro entity amount will not be accepted at the risk of application abandonment.
NOTE: If the application was previously under micro entity status, checking this box will be taken to be a notification of loss of entitlement to micro entity status.
NOTE: Checking this box will be taken to be a notification of loss of entitlement to small or micro entity status, as applicable.

NOTE: This form must be signed in accordance with 37 CFR 1.31 and 1.33. See 37 CFR 1.4 for signature requirements and certifications.

Authorized Signature _____   Date _____

Typed or printed name _____   Registration No. _____

PTOL-85 Part B (08-18) Approved for use through 01/31/2020   OMB 0651-0033   U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

DENTAL647_00017612

UNITED STATES PATENT AND TRADEMARK OFFICE

> UNITED STATES DEPARTMENT OF COMMERCE
> United States Patent and Trademark Office
> Address: COMMISSIONER FOR PATENTS
> P.O. Box 1450
> Alexandria, Virginia 22313-1450
> www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
16/031,125	07/10/2018	Philippe SALAH	N&P-51423	3171

			EXAMINER
108676	7590	12/22/2021	AZARIAN, SEYED H

Ronald M. Kachmarik
Cooper Legal Group LLC
1388 Ridge Road, Unit 1
Hinckley, OH 44233

ART UNIT	PAPER NUMBER
2667	

DATE MAILED: 12/22/2021

Determination of Patent Term Adjustment under 35 U.S.C. 154 (b)
(Applications filed on or after May 29, 2000)

The Office has discontinued providing a Patent Term Adjustment (PTA) calculation with the Notice of Allowance.

Section 1(h)(2) of the AIA Technical Corrections Act amended 35 U.S.C. 154(b)(3)(B)(i) to eliminate the requirement that the Office provide a patent term adjustment determination with the notice of allowance. See Revisions to Patent Term Adjustment, 78 Fed. Reg. 19416, 19417 (Apr. 1, 2013). Therefore, the Office is no longer providing an initial patent term adjustment determination with the notice of allowance. The Office will continue to provide a patent term adjustment determination with the Issue Notification Letter that is mailed to applicant approximately three weeks prior to the issue date of the patent, and will include the patent term adjustment on the patent. Any request for reconsideration of the patent term adjustment determination (or reinstatement of patent term adjustment) should follow the process outlined in 37 CFR 1.705.

Any questions regarding the Patent Term Extension or Adjustment determination should be directed to the Office of Patent Legal Administration at (571)-272-7702. Questions relating to issue and publication fee payments should be directed to the Customer Service Center of the Office of Patent Publication at 1-(888)-786-0101 or (571)-272-4200.

PTOL-85 (Rev. 02/11)

OMB Clearance and PRA Burden Statement for PTOL-85 Part B

The Paperwork Reduction Act (PRA) of 1995 requires Federal agencies to obtain Office of Management and Budget approval before requesting most types of information from the public. When OMB approves an agency request to collect information from the public, OMB (i) provides a valid OMB Control Number and expiration date for the agency to display on the instrument that will be used to collect the information and (ii) requires the agency to inform the public about the OMB Control Number's legal significance in accordance with 5 CFR 1320.5(b).

The information collected by PTOL-85 Part B is required by 37 CFR 1.311. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 30 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, Virginia 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, Virginia 22313-1450. Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

Privacy Act Statement

The Privacy Act of 1974 (P.L. 93-579) requires that you be given certain information in connection with your submission of the attached form related to a patent application or patent. Accordingly, pursuant to the requirements of the Act, please be advised that: (1) the general authority for the collection of this information is 35 U.S.C. 2(b)(2); (2) furnishing of the information solicited is voluntary; and (3) the principal purpose for which the information is used by the U.S. Patent and Trademark Office is to process and/or examine your submission related to a patent application or patent. If you do not furnish the requested information, the U.S. Patent and Trademark Office may not be able to process and/or examine your submission, which may result in termination of proceedings or abandonment of the application or expiration of the patent.

The information provided by you in this form will be subject to the following routine uses:

1. The information on this form will be treated confidentially to the extent allowed under the Freedom of Information Act (5 U.S.C. 552) and the Privacy Act (5 U.S.C 552a). Records from this system of records may be disclosed to the Department of Justice to determine whether disclosure of these records is required by the Freedom of Information Act.
2. A record from this system of records may be disclosed, as a routine use, in the course of presenting evidence to a court, magistrate, or administrative tribunal, including disclosures to opposing counsel in the course of settlement negotiations.
3. A record in this system of records may be disclosed, as a routine use, to a Member of Congress submitting a request involving an individual, to whom the record pertains, when the individual has requested assistance from the Member with respect to the subject matter of the record.
4. A record in this system of records may be disclosed, as a routine use, to a contractor of the Agency having need for the information in order to perform a contract. Recipients of information shall be required to comply with the requirements of the Privacy Act of 1974, as amended, pursuant to 5 U.S.C. 552a(m).
5. A record related to an International Application filed under the Patent Cooperation Treaty in this system of records may be disclosed, as a routine use, to the International Bureau of the World Intellectual Property Organization, pursuant to the Patent Cooperation Treaty.
6. A record in this system of records may be disclosed, as a routine use, to another federal agency for purposes of National Security review (35 U.S.C. 181) and for review pursuant to the Atomic Energy Act (42 U.S.C. 218(c)).
7. A record from this system of records may be disclosed, as a routine use, to the Administrator, General Services, or his/her designee, during an inspection of records conducted by GSA as part of that agency's responsibility to recommend improvements in records management practices and programs, under authority of 44 U.S.C. 2904 and 2906. Such disclosure shall be made in accordance with the GSA regulations governing inspection of records for this purpose, and any other relevant (i.e., GSA or Commerce) directive. Such disclosure shall not be used to make determinations about individuals.
8. A record from this system of records may be disclosed, as a routine use, to the public after either publication of the application pursuant to 35 U.S.C. 122(b) or issuance of a patent pursuant to 35 U.S.C. 151. Further, a record may be disclosed, subject to the limitations of 37 CFR 1.14, as a routine use, to the public if the record was filed in an application which became abandoned or in which the proceedings were terminated and which application is referenced by either a published application, an application open to public inspection or an issued patent.
9. A record from this system of records may be disclosed, as a routine use, to a Federal, State, or local law enforcement agency, if the USPTO becomes aware of a violation or potential violation of law or regulation.

Notice of Allowability	Application No. 16/031,125	Applicant(s) SALAH et al.	
	Examiner SEYED H AZARIAN	Art Unit 2667	AIA (FITF) Status Yes

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address--

All claims being allowable, PROSECUTION ON THE MERITS IS (OR REMAINS) CLOSED in this application. If not included herewith (or previously mailed), a Notice of Allowance (PTOL-85) or other appropriate communication will be mailed in due course. **THIS NOTICE OF ALLOWABILITY IS NOT A GRANT OF PATENT RIGHTS.** This application is subject to withdrawal from issue at the initiative of the Office or upon petition by the applicant. See 37 CFR 1.313 and MPEP 1308.

1. ☑ This communication is responsive to 10/12/2021.
   ☐ A declaration(s)/affidavit(s) under **37 CFR 1.130(b)** was/were filed on _____.

2. ☐ An election was made by the applicant in response to a restriction requirement set forth during the interview on _____; the restriction requirement and election have been incorporated into this action.

3. ☑ The allowed claim(s) is/are See Continuation Sheet . As a result of the allowed claim(s), you may be eligible to benefit from the **Patent Prosecution Highway** program at a participating intellectual property office for the corresponding application. For more information, please see **http://www.uspto.gov/patents/init_events/pph/index.jsp** or send an inquiry to **PPHfeedback@uspto.gov.**

4. ☑ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
   Certified copies:
   a) ☑ All   b) ☐ Some*   c) ☐ None of the:
      1. ☑ Certified copies of the priority documents have been received.
      2. ☐ Certified copies of the priority documents have been received in Application No. _____ .
      3. ☐ Copies of the certified copies of the priority documents have been received in this national stage application from the International Bureau (PCT Rule 17.2(a)).
   * Certified copies not received: _____ .

Applicant has THREE MONTHS FROM THE "MAILING DATE" of this communication to file a reply complying with the requirements noted below. Failure to timely comply will result in ABANDONMENT of this application.
THIS THREE-MONTH PERIOD IS NOT EXTENDABLE.

5. ☐ CORRECTED DRAWINGS (as "replacement sheets") must be submitted.
   ☐ including changes required by the attached Examiner's Amendment / Comment or in the Office action of Paper No./Mail Date _____ .
   Identifying indicia such as the application number (see 37 CFR 1.84(c)) should be written on the drawings in the front (not the back) of each sheet. Replacement sheet(s) should be labeled as such in the header according to 37 CFR 1.121(d).

6. ☐ DEPOSIT OF and/or INFORMATION about the deposit of BIOLOGICAL MATERIAL must be submitted. Note the attached Examiner's comment regarding REQUIREMENT FOR THE DEPOSIT OF BIOLOGICAL MATERIAL.

Attachment(s)
1. ☑ Notice of References Cited (PTO-892)
2. ☑ Information Disclosure Statements (PTO/SB/08), Paper No./Mail Date _____.
3. ☐ Examiner's Comment Regarding Requirement for Deposit of Biological Material _____.
4. ☐ Interview Summary (PTO-413), Paper No./Mail Date. _____.

5. ☐ Examiner's Amendment/Comment
6. ☑ Examiner's Statement of Reasons for Allowance
7. ☐ Other _____.

/SEYED H AZARIAN/ Primary Examiner, Art Unit 2667	

Continuation Sheet (PTOL-37)   **Application No. 16/031,125**

Continuation of 3. The allowed claim(s) is/are: 1,3-8,14,17 and 19-23

Application/Control Number: 16/031,125   Page 2
Art Unit: 2667

Notice of Pre-AIA or AIA Status

The present application, filed on or after March 16, 2013, is being examined under the

first inventor to file provisions of the AIA.

Continued Examination Under 37 CFR 1.114

A request for continued examination under 37 CFR 1.114, including the fee set forth in

37 CFR 1.17(e), was filed in this application after final rejection. Since this application is

eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e)

has been timely paid, the finality of the previous Office action has been withdrawn pursuant to

37 CFR 1.114. Applicant's submission filed on 10/12/2021 has been entered.

Response to Amendment

1.   Based on applicant's amendment, filed on 10/12/2021, see page 2 through 7 of the

remarks, with respect to cancellation of claims 2, 9, 10, 11, 12, 13, 15, 16, 18, and amended

claims 1, 17 and 22, have been fully considered and are persuasive, upon further consideration

the double patenting rejection and rejection, of 103(a) for claims 1, 3-8, 14, 17 and 19-23, are

hereby withdrawn.

The claims 1, 3-8, 14, 17 and 19-23 now renumbered as 1-14 are allowed.

REASONS FOR ALLOWANCE

2.   The following is an examiner's statement of reasons for allowance.

This invention relates generally, to a method for analyzing an image, called "analysis

image", of a dental arch of a patient, a method in which the analysis image is submitted to a deep

neural network, in order to determine at least one value of a tooth attribute

relating to a tooth represented on the analysis image.

Based on applicant's amendment, with respect to claim 1, representative of claim 17, the closest prior art of record (Kuo; Borovinskih and Kopelman), Kuo reference is directed to the field of orthodontics. More specifically, the present invention is related to methods and system for providing dynamic orthodontic assessment and treatment profiles. Borovinskih reference is directed to methods and systems for monitoring a dental patient's progress during a course of treatment. A three-dimensional model of the expected positions of the patient's teeth can be projected, in time, from a three-dimensional model of the patient's teeth prepared prior to beginning the treatment, and Kopelman reference is directed to the field of dentistry and, in particular, to a system and method for providing augmented reality enhancements for dental practitioners.

But, neither Kuo nor Borovinskih and Kopelman, teach or suggest, among other things, "acquisition, with a cellphone, by the patient, the analysis image being a photograph or an image taken from a film, and representing the dental arch of the patient; submission of the analysis image to a neural network, in order to determine at least a value of a tooth attribute relating to a tooth represented on the analysis image, steps: A) creation of a learning base comprising more than 1000 images of dental arches, or "historical images", each historical image comprising one or more zones each representing a tooth, or "historical tooth zones", to each of which, for at least one tooth attribute, a tooth attribute value is assigned; B) training of at least one deep learning device, by means of the learning base; C) submission of the analysis image to said at least one deep learning device for it to determine at least one probability relating to an attribute value of at least one tooth represented on a zone representing, at least partially, said tooth in the analysis image, or "analysis tooth zone".

These key features in combination with the other features of the claimed invention are neither taught nor suggested by (Kuo; Borovinskih and Kopelman) prior art of record.

Any comments considered necessary by applicant must be submitted no later than the payment of the issue fee and, to avoid processing delays, should preferably accompany the issue

Application/Control Number: 16/031,125   Page 4
Art Unit: 2667

fee. Such submissions should be clearly labeled "Comments on Statement of Reasons for

Allowance."

Contact Information

3.   Any inquiry concerning this communication or earlier communications from the

examiner should be directed to Seyed Azarian whose telephone number is (571) 272-7443. The

examiner can normally be reached on Monday through Thursday from 6:00 a.m. to 7:30 p.m.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's

supervisor, Matthew Bella, can be reached at (571) 272-7778. The fax phone number for the

organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent

Application information Retrieval (PAIR) system. Status information for published application

may be obtained from either Private PAIR or Public PAIR.

Status information about the PAIR system, see http:// pair-direct.uspto.gov. Should you have

questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at

866-217-9197 (toll-free).
/SEYED H AZARIAN/
Primary Examiner, Art Unit 2667
December 6, 2021

EXHIBIT A-2

deep learning   ✕   🔍

Dictionary	Thesaurus

deep learning *noun*

computing

: a form of machine learning in which the computer network rapidly teaches itself to understand a concept without human intervention by performing a large number of iterative calculations on an extremely large dataset

Deep learning involves feeding machines lots of data so the AI can learn patterns itself without requiring humans to program knowledge into the machine.

— Mike Cherney

In *deep learning*, multiple neural networks are "stacked" on top of each other, or layered, in order to create models that are even better at prediction because each new layer learns from the ones before it.

— Derrick Harris

→ sometimes used before another noun

The purpose of many *deep learning* systems is to minimize the amount of time spent feeding or quizzing a piece of software …

— Erik Sofge

Recent Examples on the Web

In particular, cardiologist Eric Topol delivered a compelling overview of how *deep learning* has been shown in research to detect seemingly unrelated diseases from X-rays, cardiograms, and retinal scans.

— Benj Edwards, *Ars Technica*, 18 Oct. 2023

Stable Diffusion is a *deep learning*, text-to-image model first released last year by the startup Stability AI, and its publicly available code has been repurposed and modified in ways that violate its user agreement.

— Miles Klee, *Rolling Stone*, 17 Oct. 2023

Machine learning researchers at Stanford University used a *deep learning* model to analyze wearable activity and sleep data from pregnant participants.

— Katie Palmer, *STAT*, 28 Sep. 2023

See More ⌄

These examples are programmatically compiled from various online sources to illustrate current usage of the word 'deep learning.' Any opinions expressed in the examples do not represent those of Merriam-Webster or its editors. Send us feedback about these examples.

First Known Use

Dictionary   Thesaurus

The first known use of *deep learning* **was in 1986**

See more words from the same year

deep kiss
deep learning
deepmost

See More Nearby Entries ›

Style   MLA

"Deep learning." *Merriam-Webster.com Dictionary*, Merriam-Webster, https://www.merriam-webster.com/dictionary/deep%20learning. Accessed 5 Nov. 2023.

Copy Citation

Facebook   Twitter

Last Updated: 27 Oct 2023 - Updated example sentences

Love words? Need even more definitions?

Subscribe to America's largest dictionary and get thousands more definitions and advanced search—ad free!

MERRIAM-WEBSTER UNABRIDGED
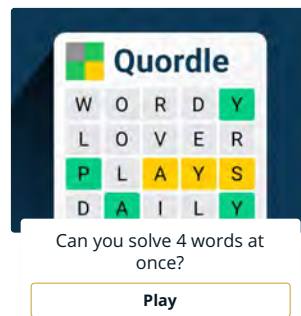
WORD OF THE DAY

Dictionary

Thesaurus

Get Word of the Day daily email!

Your email address   SUBSCRIBE

Games & Quizzes

Quordle
Can you solve 4 words at once?

Play

Blossom Word Game
You can make only 12 words. Pick the best ones!

Play

Missing Letter
A crossword with a twist

Play

Spelling Bee Quiz
Can you outdo past winners of the National Spelli...

Take the quiz

OTHER MERRIAM-WEBSTER DICTIONARIES

MERRIAM-WEBSTER'S UNABRIDGED DICTIONARY

SCRABBLE® WORD FINDER

Dictionary

Thesaurus

Browse the Dictionary: A B C D E F G H I J K L M N O P Q R S T U V W X Y Z 0-9 BIO GEO

Home | Help | About Us | Shop | Advertising Info | Dictionary API | Contact Us | Join MWU | Videos | Word of the Year |
Kid's Dictionary | Law Dictionary | Medical Dictionary | Privacy Policy | Terms of Use

Browse the Thesaurus | Browse the Medical Dictionary | Browse the Legal Dictionary | Browse the Kid's Dictionary

© 2023 Merriam-Webster, Incorporated

Information from your device can be used to personalize your ad experience.

Do not sell or share my personal information.

EXHIBIT A-3

Dictionary.com   Thesaurus.com

☰   Start typing any word or phrase   🔍

deep learning

[**deep**-lur-ning]   SHOW IPA   🔊   ☆

noun Computers.
1. an advanced type of machine learning that uses multilayered neural networks to establish nested hierarchical models for data processing and analysis, as in image recognition or natural language processing, with the goal of self-directed information processing.

QUESTION 1 OF 11

Which is a synonym for SEPIA?

carmine

| sorrel |

| canary |

TAKE THE QUIZ TO FIND OUT

ORIGIN OF DEEP LEARNING[1]

First recorded in 1985–90

See also cognitive computing.

WORDS NEARBY DEEP LEARNING

deep fryer, deep frying, deep green, deep-kiss, deep-laid, deep learning, deep link, deep linking, deep-litter, deeply, deep mourning

DICTIONARY.COM UNABRIDGED
BASED ON THE RANDOM HOUSE UNABRIDGED DICTIONARY, © RANDOM HOUSE, INC. 2023

HOW TO USE DEEP LEARNING IN A SENTENCE

They found that researchers, driven by the exploding data requirements of *deep learning*, gradually abandoned asking for people's consent.
THIS IS HOW WE LOST CONTROL OF OUR FACES | KAREN HAO | FEBRUARY 5, 2021 | MIT TECHNOLOGY REVIEW

In the last few years, research has shown that *deep learning* can match expert-level performance in medical imaging tasks like early cancer detection and eye disease diagnosis.
AI COULD MAKE HEALTHCARE FAIRER—BY HELPING US BELIEVE WHAT PATIENTS SAY | KAREN HAO | JANUARY 22, 2021 | MIT TECHNOLOGY REVIEW

The prevailing wisdom in *deep learning* research is that the more data you throw at an algorithm, the better it will learn.
HOW MIRRORING THE ARCHITECTURE OF THE HUMAN BRAIN IS SPEEDING UP AI LEARNING | EDD GENT | JANUARY 18, 2021 | SINGULARITY HUB

Advanced technologies such as *deep learning* algorithms are also playing an increasingly critical role in the development of quantum computing research.
THESE FIVE AI DEVELOPMENTS WILL SHAPE 2021 AND BEYOND | JASON SPARAPANI | JANUARY 14, 2021 | MIT TECHNOLOGY REVIEW

Some experts have placed their bets on neurosymbolic AI, which combines *deep learning* with symbolic knowledge systems.
FIVE WAYS TO MAKE AI A GREATER FORCE FOR GOOD IN 2021 | KAREN HAO | JANUARY 8, 2021 | MIT TECHNOLOGY REVIEW

SEE MORE EXAMPLES
⌄

WORD OF THE DAY
November 05, 2023

zeitgeber

noun | [**tsahyt**-gey-ber] 🔊

SEE DEFINITION

‹ ›

Sponsored

Browse

Aa Bb Cc Dd Ee Ff Gg Hh Ii Jj Kk Ll Mm Nn Oo Pp Qq Rr Ss Tt Uu Vv Ww Xx Yy Zz

Shop

Contact us

Advertise with us

Cookies, terms, & privacy

Do Not Sell My Info

Get the Word of the Day every day!

Enter your email address   **Sign up**

My account

© 2023 Dictionary.com, LLC

EXHIBIT A-4

Hewlett Packard Enterprise

Deep Learning

What is Deep Learning?

Deep learning is a type of machine learning that uses algorithms meant to function in a manner similar to the human brain.

Related to AI and machine learning

Deep learning is a subset of machine learning (ML), which is itself a subset of artificial intelligence (AI). The concept of AI has been around since the 1950s, with the goal of making computers able to think and reason in a way similar to humans. As part of making machines able to think, ML is focused on how to make them learn without being explicitly programmed. Deep learning goes beyond ML by creating more complex hierarchical models meant to mimic how humans learn new information.

Related HPE Solutions, Products, or Services

Deep Learning and Machine Learning Solutions

HPE and NVIDIA Deep Learning Collaboration

Neural networks drive deep learning

In the context of AI and ML, a model is a mathematical algorithm that is trained to come to the same result or prediction that a human expert would when provided the same information. In deep learning, the algorithms are inspired by the structure of the human brain and known as neural networks. These neural networks are built from interconnected network switches designed to learn to recognize patterns in the same way the human brain and nervous system does.

Related Topics

Artificial Intelligence

Machine Learning

Deep learning driving the future

Many recent advances in AI were made possible by deep learning. From recommendations on streaming services to voice assistant technologies to autonomous driving, the ability to identify patterns and classify many different types of information is crucial for processing vast amounts of data with little to no human input.

How does deep learning work?

While the original goal for AI was broadly to make machines able to do things that would otherwise require human intelligence, the idea has been refined in the decades since. Francois Chollet, AI researcher at Google and creator of the machine-learning software library Keras, says, "Intelligence is not skill itself, it's not what you can do, it's how well and how efficiently you can learn new things."[1]

Deep learning is focused on improving that process of having machines learn new things. With rule-based AI and ML, a data scientist determines the rules and data set features to include in models, which drives how those models operate. With deep learning, the data scientist feeds raw data into an algorithm. The system then analyzes that data, without specific rules or features preprogrammed into it. Once the system makes its predictions, they are checked against a separate set of data for accuracy. The level of accuracy of these predictions—or lack thereof—then informs the next set of predictions the system makes.

The "deep" in deep learning refers to the many layers the neural network accumulates over time, with performance improving as the network gets deeper. Each level of the network processes its input data in a specific way, which then informs the next layer. So the output from one layer becomes the input for the next.

Training deep learning networks is time consuming and requires large amounts of data to be ingested and tested against as the system gradually refines its model. Neural nets have been around since the 1950s, but only in recent years have both computational power and data storage capabilities advanced to the point where deep learning algorithms can be used to create exciting new technologies. For example, deep learning neural networks that have made it possible for computers to carry out tasks like speech recognition, computer vision, bioinformatics, and medical image analysis.

1. Lex Fridman Podcast #120, " François Chollet: Measures of Intelligence," August 2020.

Deep learning vs. machine learning

While all deep learning is machine learning, not all machine learning is deep learning. Both technologies involve training against test data to determine which model best fits the data. However, traditional machine learning methods require a certain level of human interaction to preprocess the data before the algorithms can be applied.

Machine learning is a subset of artificial intelligence. Its aim is to give computers the ability to learn without being specifically programmed on what output to deliver. The algorithms used by machine learning help the computer learn how to recognize things. This training can be tedious and require a significant amount of human effort.

Deep learning algorithms go a step further by creating hierarchical models meant to mirror our own brain's thought processes. It uses a multi-layered neural network that does not require preprocessing the input data in order to produce a result. Data scientists feed the raw data into the algorithm, the system analyzes the data based on what it already knows and what it can infer from the new data, and makes a prediction.

The advantage of deep learning is that it can process data in ways that simple rules-based AI cannot. The technology can be used to drive clear business outcomes as diverse as improved fraud detection, increased crop yields, improved accuracy of warehouse inventory control systems, and many others.

Current applications of deep learning

Companies in many sectors are applying deep learning models to address a variety of use cases. Below are just a few of the many applications of deep learning in the real world.

Healthcare: Today's medical industry is generating vast amounts of data. Being able to quickly and accurately analyze this data can contribute to improved patient outcomes in a number of ways. Deep learning algorithms are being applied in areas such as medical research, imaging analytics, disease prevention, guided drug development, and natural language processing—which can be especially helpful for filling in free text clinical notes in electronic health records (EHRs).

Manufacturing: Manufacturers need to deliver higher quality products and services faster and with lower costs. Many companies are adopting computer-aided engineering (CAE) to reduce the time, expense, and materials needed to develop physical prototypes to test new products. Deep learning can be used to model very complex patterns in multidimensional data and improve the analytics accuracy of testing data.

Financial services: Fraud is a growing problem in many industries, but particularly so for financial service providers. Deep learning can be used to identify out-of-pattern behavior quickly and cost effectively. Insights delivered from deep learning models can also help more accurately evaluate the credit risk of a loan applicant, predict stock values, automate back-office operations, and advise clients on financial products.

Public sector: As more departments, systems, and processes become digitized, government agencies can use deep learning to increase automation and make civil servants more efficient. Image detection and classification can make it easier for law enforcement to find persons of interest in public spaces. Visa and immigration applications can be streamlined with algorithms to automate certain aspects of processing. Airports are using deep learning to improve security, enhance operations, and automate queue management. Deep learning models can even be used to help predict traffic conditions and allow local authorities to take proactive steps to ease road congestion.

Speed deep learning adoption with HPE

Deep learning is now more accessible than ever before for organizations of any size. From getting started to optimizing to scaling, HPE can guide your journey to accelerate data insights that lead to breakthrough innovations. We have the infrastructure building blocks, expertise, and access to validated partners to meet your business goals.

Unravel the complexity of deep learning and create your ideal solution. HPE's industry-leading high-performance compute, intelligent data platforms, and high-speed networking fabric allow you to deploy deep learning at any scale. Rapidly move beyond proofs-of-concept with **HPE Pointnext Advisory and Professional Services**—providing the expertise and services to accelerate your AI project deployments from years to months to weeks. And gain instant access to the AI tools and data you need using **HPE Ezmeral ML Ops**, a container-based solution to support every stage of the machine learning lifecycle.

The **HPE Deep Learning Cookbook** provides a set of tools to characterize deep learning workloads and recommend the optimal hardware/software stack for any given workload. The recommendations presented through our Deep Learning Cookbook are based on a massive collection of performance results for various deep learning workloads on different technology stacks and analytical performance models. The combination of real measurements and analytical performance models enables us to estimate the performance of any workload and to recommend an optimal stack for that workload.

Together with NVIDIA, HPE offers a leading portfolio of optimized AI and deep learning solutions. We enable deep learning through online and instructor-led workshops, reference architectures, and benchmarks on NVIDIA GPU accelerated applications. Our solutions are differentiated by proven expertise, the largest deep learning ecosystem, and AI software frameworks.

How can we help?

Help me find something

How to Buy

Product Support

Email Sales

Chat with Sales

COMPANY ⌄

About HPE

Accessibility

Careers

Contact Us

Corporate Responsibility

Global Diversity & Inclusion

HPE Modern Slavery Transparency Statement (PDF)

Hewlett Packard Labs

Investor Relations

Leadership

Public Policy

LEARN ABOUT ⌄

Artificial Intelligence

Cloud Computing

Containers

Machine Learning

Enterprise Glossary

NEWS AND EVENTS ⌃

Newsroom

HPE Discover

Events

Webinars

PARTNERS ⌃

Partner Ready program

Partner Ready Vantage program

Find a Partner

Certifications

HPE GreenLake Marketplace

SUPPORT ⌃

Product Support

Software & Drivers

Warranty Check

Enhanced Support Services

Education and Training

Product Return and Recycling

OEM Solutions

COMMUNITIES ⌃

HPE Community

Aruba Airheads

HPE Tech Pro Community

HPE Developer

All Blogs and Forums

CUSTOMER RESOURCES   ⌃

Customer Stories

How To Buy

Financial Services

HPE Customer Centers

Email Signup

HPE MyAccount

Resource Library

Video Gallery

Voice of the Customer Signup

🌐   UNITED STATES (EN)

© Copyright 2023 Hewlett Packard Enterprise Development LP

Privacy | Terms of Use | Ad Choices & Cookies | Do Not Sell or Share My Personal Information | Sitemap

EXHIBIT A-5

What is deep learning?

Deep learning enables systems to cluster data and make predictions with incredible accuracy

Learn about watsonx.ai   →

What is deep learning?

Your privacy choices ✓✗

Let's talk

Deep learning vs. machine learning

How deep learning works

Deep learning applications

Deep learning hardware requirements

Related solutions

Resources

Take the next step

What is deep learning?

Deep learning is a subset of machine learning, which is essentially a neural network with three or more layers. These neural networks attempt to simulate the behavior of the human brain—albeit far from matching its ability—allowing it to "learn" from large amounts of data. While a neural network with a single layer can still make approximate predictions, additional hidden layers can help to optimize and refine for accuracy.

Deep learning drives many artificial intelligence (AI) applications and services that improve automation, performing analytical and physical tasks without human intervention. Deep learning technology lies behind everyday products and services (such as digital assistants, voice-enabled TV remotes, and credit card fraud detection) as well as emerging technologies (such as self-driving cars).

Now available: watsonx.ai

The all new enterprise studio that brings together traditional machine learning along with new generative AI capabilities powered by foundation models

Your privacy choices

Let's talk

Try watsonx.ai →

Begin your journey to AI

Learn how to scale AI →

Explore the AI Academy →

Deep learning vs. machine learning

If deep learning is a subset of machine learning, how do they differ? Deep learning distinguishes itself from classical machine learning by the type of data that it works with and the methods in which it learns.

Machine learning algorithms leverage structured, labeled data to make predictions—meaning that specific features are defined from the input data for the model and organized into tables. This doesn't necessarily mean that it doesn't use unstructured data; it just means that if it does, it generally goes through some pre-processing to organize it into a structured format.

Deep learning eliminates some of data pre-processing that is typically involved with machine learning. These algorithms can ingest and process unstructured data, like text and images, and it automates feature extraction, removing some of the dependency on human experts. For example, let's say that we had a set of photos of different pets, and we wanted to categorize by "cat", "dog", "hamster", et cetera. Deep learning algorithms can determine which features (e.g. ears) are most important to distinguish each animal from another. In machine learning, this hierarchy of features is established manually by a human expert.

Then, through the processes of gradient descent and backpropagation, the deep learning algorithm adjusts and fits itself for accuracy, allowing it to make predictions about a new photo of an animal with increased precision.

Machine learning and deep learning models are capable of different types of learning as well, which are usually categorized as supervised learning, unsupervised learning, and reinforcement learning. Supervised learning utilizes labeled datasets to categorize or make predictions; this requires some kind of human intervention to label input data correctly. In contrast, unsupervised learning doesn't require labeled datasets, and instead, it detects patterns in the data, clustering them by any distinguishing characteristics. Reinforcement learning is a process in which a model learns to become

Your privacy choices ✓✗

Let's talk

more accurate for performing an action in an environment based on feedback in order to maximize the reward.

For a deeper dive on the nuanced differences between the different technologies, see "AI vs. Machine Learning vs. Deep Learning vs. Neural Networks: What's the Difference?"

For a closer look at the specific differences between supervised and unsupervised learning, see "Supervised vs. Unsupervised Learning: What's the Difference?"

How deep learning works

Deep learning neural networks, or artificial neural networks, attempts to mimic the human brain through a combination of data inputs, weights, and bias. These elements work together to accurately recognize, classify, and describe objects within the data.

Deep neural networks consist of multiple layers of interconnected nodes, each building upon the previous layer to refine and optimize the prediction or categorization. This progression of computations through the network is called forward propagation. The input and output layers of a deep neural network are called *visible* layers. The input layer is where the deep learning model ingests the data for processing, and the output layer is where the final prediction or classification is made.

Another process called backpropagation uses algorithms, like gradient descent, to calculate errors in predictions and then adjusts the weights and biases of the function by moving backwards through the layers in an effort to train the model. Together, forward propagation and backpropagation allow a neural network to make predictions and correct for any errors accordingly. Over time, the algorithm becomes gradually more accurate.

The above describes the simplest type of deep neural network in the simplest terms. However, deep learning algorithms are incredibly complex, and there are different types of neural networks to address specific problems or datasets. For example,

– *Convolutional neural networks (CNNs),* used primarily in computer vision and image classification applications, can detect features and patterns within an image, enabling tasks, like object detection or recognition. In 2015, a CNN bested a human in an object recognition challenge for the first time.
– *Recurrent neural network (RNNs)* are typically used in natural language and speech recognition applications as it leverages sequential or times series data.

Deep learning applications

Your privacy choices ✓✕

Let's talk

Real-world deep learning applications are a part of our daily lives, but in most cases, they are so well-integrated into products and services that users are unaware of the complex data processing that is taking place in the background. Some of these examples include the following:

Law enforcement

Deep learning algorithms can analyze and learn from transactional data to identify dangerous patterns that indicate possible fraudulent or criminal activity. Speech recognition, computer vision, and other deep learning applications can improve the efficiency and effectiveness of investigative analysis by extracting patterns and evidence from sound and video recordings, images, and documents, which helps law enforcement analyze large amounts of data more quickly and accurately.

Financial services

Financial institutions regularly use predictive analytics to drive algorithmic trading of stocks, assess business risks for loan approvals, detect fraud, and help manage credit and investment portfolios for clients.

Customer service

Many organizations incorporate deep learning technology into their customer service processes. Chatbots—used in a variety of applications, services, and customer service portals—are a straightforward form of AI. Traditional chatbots use natural language and even visual recognition, commonly found in call center-like menus. However, more sophisticated chatbot solutions attempt to determine, through learning, if there are multiple responses to ambiguous questions. Based on the responses it receives, the chatbot then tries to answer these questions directly or route the conversation to a human user.

Virtual assistants like Apple's Siri, Amazon Alexa, or Google Assistant extends the idea of a chatbot by enabling speech recognition functionality. This creates a new method to engage users in a personalized way.

Healthcare

The healthcare industry has benefited greatly from deep learning capabilities ever since the digitization of hospital records and images. Image recognition applications can support medical imaging specialists and radiologists, helping them analyze and assess more images in less time.

Your privacy choices ☑☒

Let's talk

Deep learning hardware requirements

Deep learning requires a tremendous amount of computing power. High performance *graphical processing units (GPUs)* are ideal because they can handle a large volume of calculations in multiple cores with copious memory available. However, managing multiple GPUs on-premises can create a large demand on internal resources and be incredibly costly to scale.

Related solutions

watsonx

IBM watsonx is a portfolio of business-ready tools, applications and solutions, designed to reduce the costs and hurdles of AI adoption while optimizing outcomes and responsible use of AI.

Explore IBM watsonx →

watsonx Assistant - AI Chatbot

watsonx Assistant is the AI chatbot for business. This enterprise artificial intelligence technology enables users to build conversational AI solutions.

Explore watsonx Assistant →

IBM Watsonx Studio

Build, run and manage AI models. Prepare data and build models on any cloud using open source code or visual modeling. Predict and optimize your outcomes.

Explore Watsonx Studio →

Your privacy choices

Let's talk

Resources

How-to	Artificial Intelligence ebook	Article	Article
Free, hands-on learning for generative AI technologies		**An introduction to deep learning**	**Deep learning architectures**
Learn the fundamental concepts for AI and generative AI, including prompt engineering, large language models and the best open source projects.		Explore this branch of machine learning that's trained on large amounts of data and deals with computational units working in tandem to perform predictions	The rise of artificial intelligence
Learn more →	Download the ebook →	Go deeper with IBM Developer →	Learn more about architectures →

Take the next step

For decades now, IBM has been a pioneer in the development of AI technologies and deep learning, highlighted by the development of IBM watsonx Assistant, IBM's AI chatbot. One of the earliest accomplishments in deep learning technology, watsonx Assistant is now a trusted solution for enterprises looking to apply advanced natural language processing and machine learning techniques to their systems using a proven tiered approach to AI adoption and implementation.

Get started with IBM Watsonx® Studio today	→

Your privacy choices ✓✗

Let's talk

EXHIBIT A-6

What is Deep Learning?

Deep learning defined

Deep learning is a subset of machine learning (ML), where artificial neural networks—algorithms modeled to work like the human brain—learn from large amounts of data.

How does deep learning work?

Deep learning is powered by layers of neural networks, which are algorithms loosely modeled on the way human brains work. Training with large amounts of data is what configures the neurons in the neural network. The result is a deep learning model which, once trained, processes new data. Deep learning models take in information from multiple datasources and analyze that data in real time, without the need for human intervention. In deep learning, graphics processing units (GPUs) are optimized for training models because they can process multiple computations simultaneously.

Build deep learning and machine learning models

Deep learning is what drives many artificial intelligence (AI) technologies that can improve automation and analytical tasks. Most people encounter deep learning every day when they browse the internet or use their mobile phones. Among countless other applications, deep learning is used to generate captions for YouTube videos, performs speech recognition on phones and smart speakers, provides facial recognition for photographs, and enables self-driving cars. And as data scientists and researchers tackle increasingly complex deep learning projects—leveraging deep learning frameworks—this type of artificial intelligence will only become a bigger part of our daily lives.

What is the difference between deep learning and neural networks?

Deep learning vs. neural networks

In simple terms, deep learning is a name for neural networks with many layers.

To make sense of observational data, such as photos or audio, neural networks pass data through interconnected layers of nodes. When information passes through a layer, each node in that layer performs simple operations on the data and selectively passes the results to other nodes. Each subsequent layer focuses on a higher-level feature than the last, until the network creates the output.

In between the input layer and the output layer are hidden layers. This is where the distinction comes in between neural networks and deep learning: A basic neural network might have one or two hidden layers, while a deep learning network might have dozens—or even hundreds—of layers. Increasing the number of different layers and nodes may increase the accuracy of a network. However, more layers can also mean that a model will require more parameters and computational resources.

Deep learning classifies information through layers of neural networks, which have a set of inputs that receive raw data. For example, if a neural network is trained with images of birds, it can be used to recognize images of birds. More layers enable more precise results, such as distinguishing a crow from a raven as compared to distinguishing a crow from a chicken. Deep neural networks, which are behind deep learning algorithms, have several hidden layers between the input and output nodes—which means that they are able to accomplish more complex data classifications. A deep learning algorithm must be trained with large sets of data, and the more data it receives, the more accurate it will be; it will need to be fed thousands of pictures of birds before it is able to accurately classify new pictures of birds.

When it comes to neural networks, training the deep learning model is very resource intensive. This is when the neural network ingests inputs, which are processed in hidden layers using weights (parameters that represent the strength of the connection between the inputs) that are adjusted during training, and the model then puts out a prediction. Weights are adjusted based on training inputs in order to make better predictions. Deep learning models spend a lot of time in training large amounts of data, which is why high-performance compute is so important.

GPUs are optimized for data computations, and are designed for speedy performance of large-scale matrix calculations. GPUs are best suited for parallel execution for large scale machine learning (ML) and deep learning problems. As a result, ML applications that perform high numbers of computations on large amounts of structured or unstructured data—such as image, text, and video—enjoy good performance.

Drive real-time decisions with deep learning on Exadata (0:23)

Top 5 reasons to use deep learning

One major benefit of deep learning is that its neural networks are used to reveal hidden insights and relationships from data that were previously not visible. With more robust machine learning models that can analyze large, complex data, companies can improve fraud detection, supply chain management, and cybersecurity by leveraging the following:

Analyzing unstructured data

Deep learning algorithms can be trained to look at text data by analyzing social media posts, news, and surveys to provide valuable business and custom

Data labeling

Deep learning requires labeled data for training. Once trained, it can label new data and identify different types of data on its own.

Feature engineering

A deep learning algorithm can save time because it does not require humans to extract features manually from raw data.

Efficiency

When a deep learning algorithm is properly trained, it can perform thousands of tasks over and over again, faster than humans.

Training

The neural networks used in deep learning have the ability to be applied to many different data types and applications. Additionally, a deep learning model can adapt by retraining it with new data.

What's the difference between AI, machine learning, and deep learning?

AI, machine learning, and deep learning are all related, but they have distinct features:

Artificial intelligence (AI)

Artificial intelligence allows computers, machines, or robots to mimic the capabilities of a human, such as making decisions, recognizing objects, solving problems, and understanding language.

Machine learning (ML)

Machine learning is a subset of AI centered on building applications that can learn from data to improve their accuracy over time, without human intervention. Machine learning algorithms can be trained to find patterns to make better decisions and predictions, but this typically requires human intervention.

Deep learning

Deep learning is a subset of machine learning that enables computers to solve more complex problems. Deep learning models are also able to create new features on their own.

Discover the differences between AI, machine learning, and deep learning

5 uses for deep learning

Social media

Deep learning can be used to analyze a large number of images, which can help social networks find out more about their users. This improves targeted ads and follow suggestions.

Finance

Neural networks in deep learning can be used to predict stock values and develop trading strategies, and can also spot security threats and protect against fraud.

Healthcare

Deep learning can play a pivotal role in the field of healthcare by analyzing trends and behaviors to predict illnesses in patients. Healthcare workers can also employ deep learning algorithms to decide the optimal tests and treatments for their patients.

Cybersecurity

Deep learning can detect advanced threats better than traditional malware solutions by recognizing new, suspicious activities rather than responding to a database of known threats.

Digital assistants

Digital assistants represent some of the most common examples of deep learning. With the help of natural language processing (NLP), Siri, Cortana, Google, and Alexa can respond to questions and adapt to user habits.

Roadblocks to applying deep learning

While new uses for deep learning are being uncovered, it is still an evolving field with certain limitations:

Large amounts of data

In order to achieve more insightful and abstract answers, deep learning requires large amounts of data to train on. Similar to a human brain, a deep learning algorithm needs examples so that it can learn from mistakes and improve its outcome.

Lack of flexibility

Machines are still learning in very narrow ways, which can lead to mistakes. Deep learning networks need data to solve a specific problem. If asked to perform a task outside of that scope, it will most likely fail.

Lack of transparency

While it sifts through millions of data points to find patterns, it can be difficult to understand how a neural network arrives at its solution. This lack of tra[...] data makes it difficult to identify undesired biases and explain predictions.

Despite these hurdles, data scientists are getting closer and closer to building highly accurate deep learning models that can learn without supervision—[...] faster and less labor intensive.

Call US Sales

+1.800.633.0738

Complete list of local country numbers

Deep learning products and solutions

With the explosion of business data, data scientists need to be able to explore and build deep learning models quickly and with more flexibility than traditional on-premises IT hardware can provide.

Oracle Cloud Infrastructure (OCI) offers the best price-performance compute for data-intensive workloads, fast cloud storage, and low-latency, high-throughput networking with 100 Gbps RDMA. OCI also provides GPU compute instances for deep learning, easy-to-deploy images, and the flexibility to run a single-GPU workstation or cluster of multi-GPU shapes.

For building, training, and deploying machine learning models on high-performance cloud infrastructure, try Oracle Cloud Infrastructure Data Science. Data scientists can build and train deep learning models in much less time using NVIDIA GPUs in notebook sessions. They can also select the amount of compute and storage resources they need to tackle projects of any size without worrying about provisioning or maintaining infrastructure. On top of that, OCI Data Science accelerates model building by streamlining data science tasks, such as data access, algorithm selection, and model explanation.

Discover: High-performance computing solutions on OCI

Try deep learning on Oracle Cloud

☁

Oracle Cloud Free Tier

Build, test, and deploy applications on Oracle Cloud—for free.

Sign up now

Deep learning topics

Deep learning defined
How does deep learning work?
What is the difference between deep learning and **neur**al networks?
Top 5 reasons to use deep learning
What's the difference between AI, machine learning, and deep learning?
5 uses for deep learning
Roadblocks to applying deep learning
Deep learning products and solutions

Call US Sales

+1.800.633.0738

Complete list of local country numbers

EXHIBIT A-7

SSD: Single Shot MultiBox Detector

Wei Liu[1], Dragomir Anguelov[2], Dumitru Erhan[3], Christian Szegedy[3],
Scott Reed[4], Cheng-Yang Fu[1], Alexander C. Berg[1]

[1]UNC Chapel Hill [2]Zoox Inc. [3]Google Inc. [4]University of Michigan, Ann-Arbor
[1]wliu@cs.unc.edu, [2]drago@zoox.com, [3]{dumitru,szegedy}@google.com,
[4]reedscot@umich.edu, [1]{cyfu,aberg}@cs.unc.edu

Abstract. We present a method for detecting objects in images using a single deep neural network. Our approach, named SSD, discretizes the output space of bounding boxes into a set of default boxes over different aspect ratios and scales per feature map location. At prediction time, the network generates scores for the presence of each object category in each default box and produces adjustments to the box to better match the object shape. Additionally, the network combines predictions from multiple feature maps with different resolutions to naturally handle objects of various sizes. SSD is simple relative to methods that require object proposals because it completely eliminates proposal generation and subsequent pixel or feature resampling stages and encapsulates all computation in a single network. This makes SSD easy to train and straightforward to integrate into systems that require a detection component. Experimental results on the PASCAL VOC, COCO, and ILSVRC datasets confirm that SSD has competitive accuracy to methods that utilize an additional object proposal step and is much faster, while providing a unified framework for both training and inference. For 300×300 input, SSD achieves 74.3% mAP[1] on VOC2007 test at 59 FPS on a Nvidia Titan X and for 512×512 input, SSD achieves 76.9% mAP, outperforming a comparable state-of-the-art Faster R-CNN model. Compared to other single stage methods, SSD has much better accuracy even with a smaller input image size. Code is available at: https://github.com/weiliu89/caffe/tree/ssd.

Keywords: Real-time Object Detection; Convolutional Neural Network

1   Introduction

Current state-of-the-art object detection systems are variants of the following approach: hypothesize bounding boxes, resample pixels or features for each box, and apply a high-quality classifier. This pipeline has prevailed on detection benchmarks since the Selective Search work [1] through the current leading results on PASCAL VOC, COCO, and ILSVRC detection all based on Faster R-CNN[2] albeit with deeper features such as [3]. While accurate, these approaches have been too computationally intensive for embedded systems and, even with high-end hardware, too slow for real-time applications.

---

[1] We achieved even better results using an improved data augmentation scheme in follow-on experiments: 77.2% mAP for 300×300 input and 79.8% mAP for 512×512 input on VOC2007. Please see Sec. 3.6 for details.

Often detection speed for these approaches is measured in seconds per frame (SPF), and even the fastest high-accuracy detector, Faster R-CNN, operates at only 7 frames per second (FPS). There have been many attempts to build faster detectors by attacking each stage of the detection pipeline (see related work in Sec. 4), but so far, significantly increased speed comes only at the cost of significantly decreased detection accuracy.

This paper presents the first deep network based object detector that does not re-sample pixels or features for bounding box hypotheses *and* and is as accurate as approaches that do. This results in a significant improvement in speed for high-accuracy detection (59 FPS with mAP 74.3% on VOC2007 test, vs. Faster R-CNN 7 FPS with mAP 73.2% or YOLO 45 FPS with mAP 63.4%). The fundamental improvement in speed comes from eliminating bounding box proposals and the subsequent pixel or feature resampling stage. We are not the first to do this (cf [4,5]), but by adding a series of improvements, we manage to increase the accuracy significantly over previous attempts. Our improvements include using a small convolutional filter to predict object categories and offsets in bounding box locations, using separate predictors (filters) for different aspect ratio detections, and applying these filters to multiple feature maps from the later stages of a network in order to perform detection at multiple scales. With these modifications—especially using multiple layers for prediction at different scales—we can achieve high-accuracy using relatively low resolution input, further increasing detection speed. While these contributions may seem small independently, we note that the resulting system improves accuracy on real-time detection for PASCAL VOC from 63.4% mAP for YOLO to 74.3% mAP for our SSD. This is a larger relative improvement in detection accuracy than that from the recent, very high-profile work on residual networks [3]. Furthermore, significantly improving the speed of high-quality detection can broaden the range of settings where computer vision is useful.

We summarize our contributions as follows:

– We introduce SSD, a single-shot detector for multiple categories that is faster than the previous state-of-the-art for single shot detectors (YOLO), and significantly more accurate, in fact as accurate as slower techniques that perform explicit region proposals and pooling (including Faster R-CNN).
– The core of SSD is predicting category scores and box offsets for a fixed set of default bounding boxes using small convolutional filters applied to feature maps.
– To achieve high detection accuracy we produce predictions of different scales from feature maps of different scales, and explicitly separate predictions by aspect ratio.
– These design features lead to simple end-to-end training and high accuracy, even on low resolution input images, further improving the speed vs accuracy trade-off.
– Experiments include timing and accuracy analysis on models with varying input size evaluated on PASCAL VOC, COCO, and ILSVRC and are compared to a range of recent state-of-the-art approaches.

2   The Single Shot Detector (SSD)

This section describes our proposed SSD framework for detection (Sec. 2.1) and the associated training methodology (Sec. 2.2). Afterwards, Sec. 3 presents dataset-specific model details and experimental results.
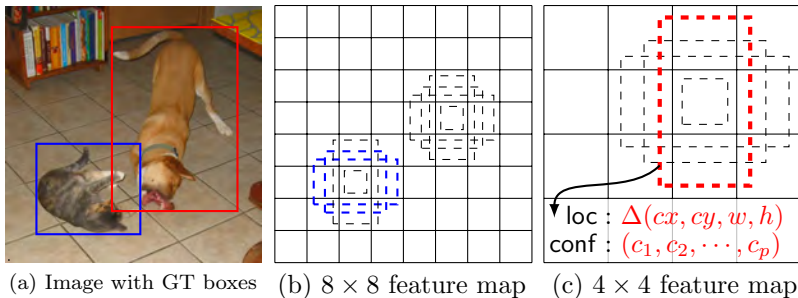
SSD: Single Shot MultiBox Detector   3

(a) Image with GT boxes   (b) 8×8 feature map   (c) 4×4 feature map

Fig. 1: **SSD framework.** (a) SSD only needs an input image and ground truth boxes for each object during training. In a convolutional fashion, we evaluate a small set (e.g. 4) of default boxes of different aspect ratios at each location in several feature maps with different scales (e.g. 8×8 and 4×4 in (b) and (c)). For each default box, we predict both the shape offsets and the confidences for all object categories (($c_1, c_2, \cdots, c_p$)). At training time, we first match these default boxes to the ground truth boxes. For example, we have matched two default boxes with the cat and one with the dog, which are treated as positives and the rest as negatives. The model loss is a weighted sum between localization loss (e.g. Smooth L1 [6]) and confidence loss (e.g. Softmax).

2.1 Model

The SSD approach is based on a feed-forward convolutional network that produces a fixed-size collection of bounding boxes and scores for the presence of object class instances in those boxes, followed by a non-maximum suppression step to produce the final detections. The early network layers are based on a standard architecture used for high quality image classification (truncated before any classification layers), which we will call the base network[2]. We then add auxiliary structure to the network to produce detections with the following key features:

Multi-scale feature maps for detection We add convolutional feature layers to the end of the truncated base network. These layers decrease in size progressively and allow predictions of detections at multiple scales. The convolutional model for predicting detections is different for each feature layer (*cf* Overfeat[4] and YOLO[5] that operate on a single scale feature map).

Convolutional predictors for detection Each added feature layer (or optionally an existing feature layer from the base network) can produce a fixed set of detection predictions using a set of convolutional filters. These are indicated on top of the SSD network architecture in Fig. 2. For a feature layer of size $m \times n$ with p channels, the basic element for predicting parameters of a potential detection is a $3 \times 3 \times p$ *small kernel* that produces either a score for a category, or a shape offset relative to the default box coordinates. At each of the $m \times n$ locations where the kernel is applied, it produces an output value. The bounding box offset output values are measured relative to a default

---

[2] We use the VGG-16 network as a base, but other networks should also produce good results.

Fig. 2: A comparison between two single shot detection models: SSD and YOLO [5]. Our SSD model adds several feature layers to the end of a base network, which predict the offsets to default boxes of different scales and aspect ratios and their associated confidences. SSD with a 300×300 input size significantly outperforms its 448×448 YOLO counterpart in accuracy on VOC2007 `test` while also improving the speed.

box position relative to each feature map location (*cf* the architecture of YOLO[5] that uses an intermediate fully connected layer instead of a convolutional filter for this step).

Default boxes and aspect ratios We associate a set of default bounding boxes with each feature map cell, for multiple feature maps at the top of the network. The default boxes tile the feature map in a convolutional manner, so that the position of each box relative to its corresponding cell is fixed. At each feature map cell, we predict the offsets relative to the default box shapes in the cell, as well as the per-class scores that indicate the presence of a class instance in each of those boxes. Specifically, for each box out of k at a given location, we compute c class scores and the 4 offsets relative to the original default box shape. This results in a total of $(c + 4)k$ filters that are applied around each location in the feature map, yielding $(c + 4)kmn$ outputs for a $m \times n$ feature map. For an illustration of default boxes, please refer to Fig. 1. Our default boxes are similar to the *anchor boxes* used in Faster R-CNN [2], however we apply them to several feature maps of different resolutions. Allowing different default box shapes in several feature maps let us efficiently discretize the space of possible output box shapes.

2.2   Training

The key difference between training SSD and training a typical detector that uses region proposals, is that ground truth information needs to be assigned to specific outputs in the fixed set of detector outputs. Some version of this is also required for training in YOLO[5] and for the region proposal stage of Faster R-CNN[2] and MultiBox[7]. Once this assignment is determined, the loss function and back propagation are applied end-to-end. Training also involves choosing the set of default boxes and scales for detection as well as the hard negative mining and data augmentation strategies.

Matching strategy During training we need to determine which default boxes correspond to a ground truth detection and train the network accordingly. For each ground truth box we are selecting from default boxes that vary over location, aspect ratio, and scale. We begin by matching each ground truth box to the default box with the best jaccard overlap (as in MultiBox [7]). Unlike MultiBox, we then match default boxes to any ground truth with jaccard overlap higher than a threshold (0.5). This simplifies the learning problem, allowing the network to predict high scores for multiple overlapping default boxes rather than requiring it to pick only the one with maximum overlap.

Training objective The SSD training objective is derived from the MultiBox objective [7,8] but is extended to handle multiple object categories. Let $x_{ij}^p = \{1, 0\}$ be an indicator for matching the i-th default box to the j-th ground truth box of category p. In the matching strategy above, we can have $\sum_i x_{ij}^p \geq 1$. The overall objective loss function is a weighted sum of the localization loss (loc) and the confidence loss (conf):

$$L(x, c, l, g) = \frac{1}{N}(L_{conf}(x, c) + \alpha L_{loc}(x, l, g)) \tag{1}$$

where N is the number of matched default boxes. If $N = 0$, wet set the loss to 0. The localization loss is a Smooth L1 loss [6] between the predicted box (l) and the ground truth box (g) parameters. Similar to Faster R-CNN [2], we regress to offsets for the center (cx, cy) of the default bounding box (d) and for its width (w) and height (h).

$$L_{loc}(x, l, g) = \sum_{i \in Pos}^{N} \sum_{m \in \{cx, cy, w, h\}} x_{ij}^k \text{smooth}_{L1}(l_i^m - \hat{g}_j^m)$$
$$\hat{g}_j^{cx} = (g_j^{cx} - d_i^{cx})/d_i^w \qquad \hat{g}_j^{cy} = (g_j^{cy} - d_i^{cy})/d_i^h \tag{2}$$
$$\hat{g}_j^w = \log\left(\frac{g_j^w}{d_i^w}\right) \qquad \hat{g}_j^h = \log\left(\frac{g_j^h}{d_i^h}\right)$$

The confidence loss is the softmax loss over multiple classes confidences (c).

$$L_{conf}(x, c) = -\sum_{i \in Pos}^{N} x_{ij}^p log(\hat{c}_i^p) - \sum_{i \in Neg} log(\hat{c}_i^0) \quad \text{where} \quad \hat{c}_i^p = \frac{\exp(c_i^p)}{\sum_p \exp(c_i^p)} \tag{3}$$

and the weight term α is set to 1 by cross validation.

Choosing scales and aspect ratios for default boxes To handle different object scales, some methods [4,9] suggest processing the image at different sizes and combining the results afterwards. However, by utilizing feature maps from several different layers in a single network for prediction we can mimic the same effect, while also sharing parameters across all object scales. Previous works [10,11] have shown that using feature maps from the lower layers can improve semantic segmentation quality because the lower layers capture more fine details of the input objects. Similarly, [12] showed that adding global context pooled from a feature map can help smooth the segmentation results.

Motivated by these methods, we use both the lower and upper feature maps for detection. Figure 1 shows two exemplar feature maps (8×8 and 4×4) which are used in the framework. In practice, we can use many more with small computational overhead.

Feature maps from different levels within a network are known to have different (empirical) receptive field sizes [13]. Fortunately, within the SSD framework, the default boxes do not necessary need to correspond to the actual receptive fields of each layer. We design the tiling of default boxes so that specific feature maps learn to be responsive to particular scales of the objects. Suppose we want to use m feature maps for prediction. The scale of the default boxes for each feature map is computed as:

$$s_k = s_{\min} + \frac{s_{\max} - s_{\min}}{m - 1}(k - 1), \quad k \in [1, m] \tag{4}$$

where $s_{\min}$ is 0.2 and $s_{\max}$ is 0.9, meaning the lowest layer has a scale of 0.2 and the highest layer has a scale of 0.9, and all layers in between are regularly spaced. We impose different aspect ratios for the default boxes, and denote them as $a_r \in \{1, 2, 3, \frac{1}{2}, \frac{1}{3}\}$. We can compute the width ($w_k^a = s_k \sqrt{a_r}$) and height ($h_k^a = s_k / \sqrt{a_r}$) for each default box. For the aspect ratio of 1, we also add a default box whose scale is $s_k' = \sqrt{s_k s_{k+1}}$, resulting in 6 default boxes per feature map location. We set the center of each default box to $(\frac{i+0.5}{|f_k|}, \frac{j+0.5}{|f_k|})$, where $|f_k|$ is the size of the k-th square feature map, $i, j \in [0, |f_k|)$. In practice, one can also design a distribution of default boxes to best fit a specific dataset. How to design the optimal tiling is an open question as well.

By combining predictions for all default boxes with different scales and aspect ratios from all locations of many feature maps, we have a diverse set of predictions, covering various input object sizes and shapes. For example, in Fig. 1, the dog is matched to a default box in the 4×4 feature map, but not to any default boxes in the 8×8 feature map. This is because those boxes have different scales and do not match the dog box, and therefore are considered as negatives during training.

Hard negative mining After the matching step, most of the default boxes are negatives, especially when the number of possible default boxes is large. This introduces a significant imbalance between the positive and negative training examples. Instead of using all the negative examples, we sort them using the highest confidence loss for each default box and pick the top ones so that the ratio between the negatives and positives is at most 3:1. We found that this leads to faster optimization and a more stable training.

Data augmentation To make the model more robust to various input object sizes and shapes, each training image is randomly sampled by one of the following options:

– Use the entire original input image.
– Sample a patch so that the *minimum* jaccard overlap with the objects is 0.1, 0.3, 0.5, 0.7, or 0.9.
– Randomly sample a patch.

The size of each sampled patch is [0.1, 1] of the original image size, and the aspect ratio is between $\frac{1}{2}$ and 2. We keep the overlapped part of the ground truth box if the center of it is in the sampled patch. After the aforementioned sampling step, each sampled patch is resized to fixed size and is horizontally flipped with probability of 0.5, in addition to applying some photo-metric distortions similar to those described in [14].

3   Experimental Results

Base network   Our experiments are all based on VGG16 [15], which is pre-trained on the ILSVRC CLS-LOC dataset [16]. Similar to DeepLab-LargeFOV [17], we convert fc6 and fc7 to convolutional layers, subsample parameters from fc6 and fc7, change pool5 from $2 \times 2 - s2$ to $3 \times 3 - s1$, and use the *à trous* algorithm [18] to fill the "holes". We remove all the dropout layers and the fc8 layer. We fine-tune the resulting model using SGD with initial learning rate 10^{-3}, 0.9 momentum, 0.0005 weight decay, and batch size 32. The learning rate decay policy is slightly different for each dataset, and we will describe details later. The full training and testing code is built on Caffe [19] and is open source at: `https://github.com/weiliu89/caffe/tree/ssd` .

3.1   PASCAL VOC2007

On this dataset, we compare against Fast R-CNN [6] and Faster R-CNN [2] on VOC2007 `test` (4952 images). All methods fine-tune on the same pre-trained VGG16 network.

Figure 2 shows the architecture details of the SSD300 model. We use conv4_3, conv7 (fc7), conv8_2, conv9_2, conv10_2, and conv11_2 to predict both location and confidences. We set default box with scale 0.1 on conv4_3[3]. We initialize the parameters for all the newly added convolutional layers with the "xavier" method [20]. For conv4_3, conv10_2 and conv11_2, we only associate 4 default boxes at each feature map location – omitting aspect ratios of $\frac{1}{3}$ and 3. For all other layers, we put 6 default boxes as described in Sec. 2.2. Since, as pointed out in [12], conv4_3 has a different feature scale compared to the other layers, we use the L2 normalization technique introduced in [12] to scale the feature norm at each location in the feature map to 20 and learn the scale during back propagation. We use the 10^{-3} learning rate for 40k iterations, then continue training for 10k iterations with 10^{-4} and 10^{-5}. When training on VOC2007 `trainval`, Table 1 shows that our low resolution SSD300 model is already more accurate than Fast R-CNN. When we train SSD on a larger 512×512 input image, it is even more accurate, surpassing Faster R-CNN by 1.7% mAP. If we train SSD with more (i.e. 07+12) data, we see that SSD300 is already better than Faster R-CNN by 1.1% and that SSD512 is 3.6% better. If we take models trained on COCO `trainval35k` as described in Sec. 3.4 and fine-tuning them on the 07+12 dataset with SSD512, we achieve the best results: 81.6% mAP.

To understand the performance of our two SSD models in more details, we used the detection analysis tool from [21]. Figure 3 shows that SSD can detect various object categories with high quality (large white area). The majority of its confident detections are correct. The recall is around 85-90%, and is much higher with "weak" (0.1 jaccard overlap) criteria. Compared to R-CNN [22], SSD has less localization error, indicating that SSD can localize objects better because it directly learns to regress the object shape and classify object categories instead of using two decoupled steps. However, SSD has more confusions with similar object categories (especially for animals), partly because we share locations for multiple categories. Figure 4 shows that SSD is very sensitive to the bounding box size. In other words, it has much worse performance on smaller

---

[3] For SSD512 model, we add extra conv12_2 for prediction, set s_{min} to 0.15, and 0.07 on conv4_3.

Method	data	mAP	aero	bike	bird	boat	bottle	bus	car	cat	chair	cow	table	dog	horse	mbike	person	plant	sheep	sofa	train	tv
Fast [6]	07	66.9	74.5	78.3	69.2	53.2	36.6	77.3	78.2	82.0	40.7	72.7	67.9	79.6	79.2	73.0	69.0	30.1	65.4	70.2	75.8	65.8
Fast [6]	07+12	70.0	77.0	78.1	69.3	59.4	38.3	81.6	78.6	86.7	42.8	78.8	68.9	84.7	82.0	76.6	69.9	31.8	70.1	74.8	80.4	70.4
Faster [2]	07	69.9	70.0	80.6	70.1	57.3	49.9	78.2	80.4	82.0	52.2	75.3	67.2	80.3	79.8	75.0	76.3	39.1	68.3	67.3	81.1	67.6
Faster [2]	07+12	73.2	76.5	79.0	70.9	65.5	52.1	83.1	84.7	86.4	52.0	81.9	65.7	84.8	84.6	77.5	76.7	38.8	73.6	73.9	83.0	72.6
Faster [2]	07+12+COCO	78.8	84.3	82.0	77.7	68.9	65.7	88.1	88.4	88.9	63.6	86.3	70.8	85.9	87.6	80.1	82.3	53.6	80.4	75.8	86.6	78.9
SSD300	07	68.0	73.4	77.5	64.1	59.0	38.9	75.2	80.8	78.5	46.0	67.8	69.2	76.6	82.1	77.0	72.5	41.2	64.2	69.1	78.0	68.5
SSD300	07+12	74.3	75.5	80.2	72.3	66.3	47.6	83.0	84.2	86.1	54.7	78.3	73.9	84.5	85.3	82.6	76.2	48.6	73.9	76.0	83.4	74.0
SSD300	07+12+COCO	79.6	80.9	86.3	79.0	**76.2**	57.6	87.3	88.2	88.6	60.5	85.4	**76.7**	**87.5**	**89.2**	84.5	81.4	55.0	81.9	**81.5**	85.9	78.9
SSD512	07	71.6	75.1	81.4	69.8	60.8	46.3	82.6	84.7	84.1	48.5	75.0	67.4	82.3	83.9	79.4	76.6	44.9	69.9	69.1	78.1	71.8
SSD512	07+12	76.8	82.4	84.7	78.4	73.8	53.2	86.2	87.5	86.0	57.8	83.1	70.2	84.9	85.2	83.9	79.7	50.3	77.9	73.9	82.5	75.3
SSD512	07+12+COCO	**81.6**	**86.6**	**88.3**	**82.4**	76.0	**66.3**	**88.6**	**88.9**	**89.1**	**65.1**	**88.4**	73.6	86.5	88.9	**85.3**	**84.6**	**59.1**	**85.0**	80.4	**87.4**	**81.2**

Table 1: **PASCAL VOC2007 `test` detection results.** Both Fast and Faster R-CNN use input images whose minimum dimension is 600. The two SSD models have exactly the same settings except that they have different input sizes (300×300 vs. 512×512). It is obvious that larger input size leads to better results, and more data always helps. Data: "07": VOC2007 `trainval`, "07+12": union of VOC2007 and VOC2012 `trainval`. "07+12+COCO": first train on COCO `trainval35k` then fine-tune on 07+12.

objects than bigger objects. This is not surprising because those small objects may not even have any information at the very top layers. Increasing the input size (e.g. from 300×300 to 512×512) can help improve detecting small objects, but there is still a lot of room to improve. On the positive side, we can clearly see that SSD performs really well on large objects. And it is very robust to different object aspect ratios because we use default boxes of various aspect ratios per feature map location.

3.2   Model analysis

To understand SSD better, we carried out controlled experiments to examine how each component affects performance. For all the experiments, we use the same settings and input size (300×300), except for specified changes to the settings or component(s).

	SSD300				
more data augmentation?		✔	✔	✔	✔
include $\{\frac{1}{2}, 2\}$ box?	✔		✔	✔	✔
include $\{\frac{1}{3}, 3\}$ box?	✔			✔	✔
use atrous?	✔	✔	✔		✔
VOC2007 `test` mAP	65.5	71.6	73.7	74.2	**74.3**

Table 2: **Effects of various design choices and components on SSD performance.**

Data augmentation is crucial. Fast and Faster R-CNN use the original image and the horizontal flip to train. We use a more extensive sampling strategy, similar to YOLO [5]. Table 2 shows that we can improve 8.8% mAP with this sampling strategy. We do not know how much our sampling strategy will benefit Fast and Faster R-CNN, but they are likely to benefit less because they use a feature pooling step during classification that is relatively robust to object translation by design.
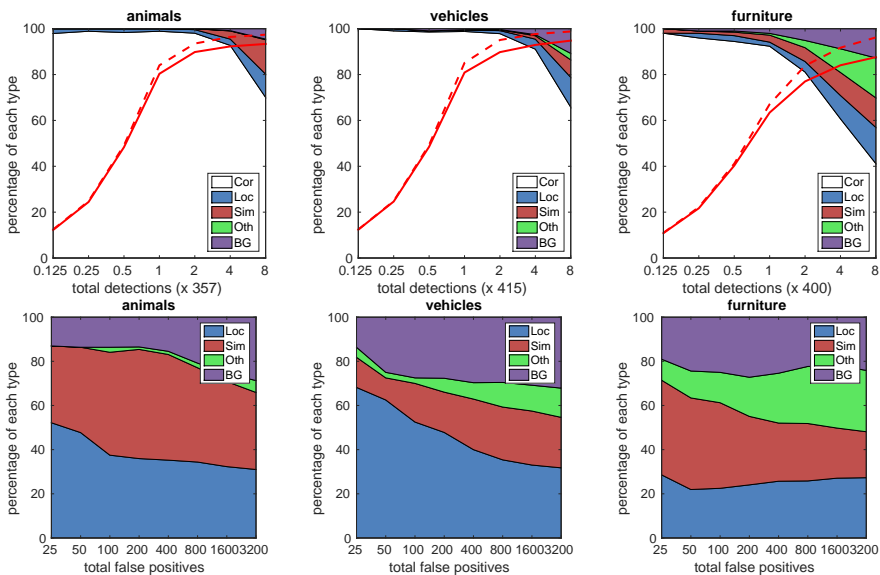
Fig. 3: **Visualization of performance for SSD512 on animals, vehicles, and furniture from VOC2007 `test`.** The top row shows the cumulative fraction of detections that are correct (Cor) or false positive due to poor localization (Loc), confusion with similar categories (Sim), with others (Oth), or with background (BG). The solid red line reflects the change of recall with strong criteria (0.5 jaccard overlap) as the number of detections increases. The dashed red line is using the weak criteria (0.1 jaccard overlap). The bottom row shows the distribution of top-ranked false positive types.
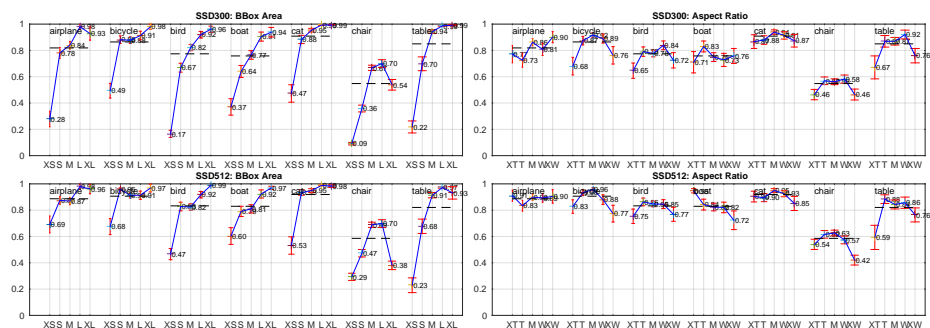
Fig. 4: **Sensitivity and impact of different object characteristics on VOC2007 `test` set using [21].** The plot on the left shows the effects of BBox Area per category, and the right plot shows the effect of Aspect Ratio. Key: BBox Area: XS=extra-small; S=small; M=medium; L=large; XL =extra-large. Aspect Ratio: XT=extra-tall/narrow; T=tall; M=medium; W=wide; XW =extra-wide.

More default box shapes is better. As described in Sec. 2.2, by default we use 6 default boxes per location. If we remove the boxes with $\frac{1}{3}$ and 3 aspect ratios, the performance drops by 0.6%. By further removing the boxes with $\frac{1}{2}$ and 2 aspect ratios, the performance drops another 2.1%. Using a variety of default box shapes seems to make the task of predicting boxes easier for the network.

Atrous is faster. As described in Sec. 3, we used the atrous version of a subsampled VGG16, following DeepLab-LargeFOV [17]. If we use the full VGG16, keeping pool5 with $2 \times 2 - s2$ and not subsampling parameters from fc6 and fc7, and add conv5_3 for prediction, the result is about the same while the speed is about 20% slower.

Prediction source layers from:						mAP use boundary boxes?		# Boxes
conv4_3	conv7	conv8_2	conv9_2	conv10_2	conv11_2	Yes	No	
✔	✔	✔	✔	✔	✔	74.3	63.4	8732
✔	✔	✔	✔	✔		**74.6**	63.1	8764
✔	✔	✔	✔			73.8	68.4	8942
✔	✔	✔				70.7	69.2	9864
✔	✔					64.2	64.4	9025
	✔					62.4	64.0	8664

Table 3: **Effects of using multiple output layers.**

Multiple output layers at different resolutions is better. A major contribution of SSD is using default boxes of different scales on different output layers. To measure the advantage gained, we progressively remove layers and compare results. For a fair comparison, every time we remove a layer, we adjust the default box tiling to keep the total number of boxes similar to the original (8732). This is done by stacking more scales of boxes on remaining layers and adjusting scales of boxes if needed. We do not exhaustively optimize the tiling for each setting. Table 3 shows a decrease in accuracy with fewer layers, dropping monotonically from 74.3 to 62.4. When we stack boxes of multiple scales on a layer, many are on the image boundary and need to be handled carefully. We tried the strategy used in Faster R-CNN [2], ignoring boxes which are on the boundary. We observe some interesting trends. For example, it hurts the performance by a large margin if we use very coarse feature maps (e.g. conv11_2 (1×1) or conv10_2 (3×3)). The reason might be that we do not have enough large boxes to cover large objects after the pruning. When we use primarily finer resolution maps, the performance starts increasing again because even after pruning a sufficient number of large boxes remains. If we only use conv7 for prediction, the performance is the worst, reinforcing the message that it is critical to spread boxes of different scales over different layers. Besides, since our predictions do not rely on ROI pooling as in [6], we do not have the *collapsing bins* problem in low-resolution feature maps [23]. The SSD architecture combines predictions from feature maps of various resolutions to achieve comparable accuracy to Faster R-CNN, while using lower resolution input images.

3.3   PASCAL VOC2012

We use the same settings as those used for our basic VOC2007 experiments above, except that we use VOC2012 `trainval` and VOC2007 `trainval` and `test` (21503 images) for training, and test on VOC2012 `test` (10991 images). We train the models with 10^{-3} learning rate for 60k iterations, then 10^{-4} for 20k iterations. Table 4 shows the results of our SSD300 and SSD512[4] model. We see the same performance trend as we observed on VOC2007 test. Our SSD300 improves accuracy over Fast/Faster R-CNN. By increasing the training and testing image size to 512×512, we are 4.5% more accurate than Faster R-CNN. Compared to YOLO, SSD is significantly more accurate, likely due to the use of convolutional default boxes from multiple feature maps and our matching strategy during training. When fine-tuned from models trained on COCO, our SSD512 achieves 80.0% mAP, which is 4.1% higher than Faster R-CNN.

Method	data	mAP	aero	bike	bird	boat	bottle	bus	car	cat	chair	cow	table	dog	horse	mbike	person	plant	sheep	sofa	train	tv
Fast[6]	07++12	68.4	82.3	78.4	70.8	52.3	38.7	77.8	71.6	89.3	44.2	73.0	55.0	87.5	80.5	80.8	72.0	35.1	68.3	65.7	80.4	64.2
Faster[2]	07++12	70.4	84.9	79.8	74.3	53.9	49.8	77.5	75.9	88.5	45.6	77.1	55.3	86.9	81.7	80.9	79.6	40.1	72.6	60.9	81.2	61.5
Faster[2]	07++12+COCO	75.9	87.4	83.6	76.8	62.9	59.6	81.9	82.0	91.3	54.9	82.6	59.0	89.0	85.5	84.7	84.1	52.2	78.9	65.5	85.4	70.2
YOLO[5]	07++12	57.9	77.0	67.2	57.7	38.3	22.7	68.3	55.9	81.4	36.2	60.8	48.5	77.2	72.3	71.3	63.5	28.9	52.2	54.8	73.9	50.8
SSD300	07++12	72.4	85.6	80.1	70.5	57.6	46.2	79.4	76.1	89.2	53.0	77.0	60.8	87.0	83.1	82.3	79.4	45.9	75.9	69.5	81.9	67.5
SSD300	07++12+COCO	77.5	90.2	83.3	76.3	63.0	53.6	83.8	82.8	92.0	59.7	82.7	63.5	89.3	87.6	85.9	84.3	52.6	82.5	74.1	88.4	74.2
SSD512	07++12	74.9	87.4	82.3	75.8	59.0	52.6	81.7	81.5	90.0	55.4	79.0	59.8	88.4	84.3	84.7	83.3	50.2	78.0	66.3	86.3	72.0
SSD512	07++12+COCO	80.0	90.7	86.8	80.5	67.8	60.8	86.3	85.5	93.5	63.2	85.7	64.4	90.9	89.0	88.9	86.8	57.2	85.1	72.8	88.4	75.9

Table 4: **PASCAL VOC2012 `test` detection results.** Fast and Faster R-CNN use images with minimum dimension 600, while the image size for YOLO is 448×448. data: "07++12": union of VOC2007 `trainval` and `test` and VOC2012 `trainval`. "07++12+COCO": first train on COCO `trainval35k` then fine-tune on 07++12.

3.4   COCO

To further validate the SSD framework, we trained our SSD300 and SSD512 architectures on the COCO dataset. Since objects in COCO tend to be smaller than PASCAL VOC, we use smaller default boxes for all layers. We follow the strategy mentioned in Sec. 2.2, but now our smallest default box has a scale of 0.15 instead of 0.2, and the scale of the default box on conv4_3 is 0.07 (e.g. 21 pixels for a 300×300 image)[5].

We use the `trainval35k` [24] for training. We first train the model with 10^{-3} learning rate for 160k iterations, and then continue training for 40k iterations with 10^{-4} and 40k iterations with 10^{-5}. Table 5 shows the results on `test-dev2015`. Similar to what we observed on the PASCAL VOC dataset, SSD300 is better than Fast R-CNN in both mAP@0.5 and mAP@[0.5:0.95]. SSD300 has a similar mAP@0.75 as ION [24] and Faster R-CNN [25], but is worse in mAP@0.5. By increasing the image size to 512×512, our SSD512 is better than Faster R-CNN [25] in both criteria. Interestingly, we observe that SSD512 is 5.3% better in mAP@0.75, but is only 1.2% better in mAP@0.5. We also observe that it has much better AP (4.8%) and AR (4.6%) for large objects, but has relatively less improvement in AP (1.3%) and AR (2.0%) for

---

[4] http://host.robots.ox.ac.uk:8080/leaderboard/displaylb.php?cls=mean&challengeid=11&compid=4

[5] For SSD512 model, we add extra conv12_2 for prediction, set $s_{\min}$ to 0.1, and 0.04 on conv4_3.

Liu et al.

Method	data	Avg. Precision, IoU:			Avg. Precision, Area:			Avg. Recall, #Dets:			Avg. Recall, Area:		
		0.5:0.95	0.5	0.75	S	M	L	1	10	100	S	M	L
Fast [6]	train	19.7	35.9	-	-	-	-	-	-	-	-	-	-
Fast [24]	train	20.5	39.9	19.4	4.1	20.0	35.8	21.3	29.5	30.1	7.3	32.1	52.0
Faster [2]	trainval	21.9	42.7	-	-	-	-	-	-	-	-	-	-
ION [24]	train	23.6	43.2	23.6	6.4	24.1	38.3	23.2	32.7	33.5	10.1	37.7	53.6
Faster [25]	trainval	24.2	45.3	23.5	7.7	26.4	37.1	23.8	34.0	34.6	12.0	38.5	54.4
SSD300	trainval35k	23.2	41.2	23.4	5.3	23.2	39.6	22.5	33.2	35.3	9.6	37.6	56.5
SSD512	trainval35k	**26.8**	**46.5**	**27.8**	**9.0**	**28.9**	**41.9**	**24.8**	**37.5**	**39.8**	**14.0**	**43.5**	**59.0**

Table 5: **COCO `test-dev2015` detection results.**

small objects. Compared to ION, the improvement in AR for large and small objects is more similar (5.4% vs. 3.9%). We conjecture that Faster R-CNN is more competitive on smaller objects with SSD because it performs two box refinement steps, in both the RPN part and in the Fast R-CNN part. In Fig. 5, we show some detection examples on COCO `test-dev` with the SSD512 model.

3.5   Preliminary ILSVRC results

We applied the same network architecture we used for COCO to the ILSVRC DET dataset [16]. We train a SSD300 model using the ILSVRC2014 DET `train` and `val1` as used in [22]. We first train the model with 10^{-3} learning rate for 320k iterations, and then continue training for 80k iterations with 10^{-4} and 40k iterations with 10^{-5}. We can achieve 43.4 mAP on the `val2` set [22]. Again, it validates that SSD is a general framework for high quality real-time detection.

3.6   Data Augmentation for Small Object Accuracy

Without a follow-up feature resampling step as in Faster R-CNN, the classification task for small objects is relatively hard for SSD, as demonstrated in our analysis (see Fig. 4). The data augmentation strategy described in Sec. 2.2 helps to improve the performance dramatically, especially on small datasets such as PASCAL VOC. The random crops generated by the strategy can be thought of as a "zoom in" operation and can generate many larger training examples. To implement a "zoom out" operation that creates more small training examples, we first randomly place an image on a canvas of $16\times$ of the original image size filled with mean values before we do any random crop operation. Because we have more training images by introducing this new "expansion" data augmentation trick, we have to double the training iterations. We have seen a consistent increase of 2%-3% mAP across multiple datasets, as shown in Table 6. In specific, Figure 6 shows that the new augmentation trick significantly improves the performance on small objects. This result underscores the importance of the data augmentation strategy for the final model accuracy.

An alternative way of improving SSD is to design a better tiling of default boxes so that its position and scale are better aligned with the receptive field of each position on a feature map. We leave this for future work.
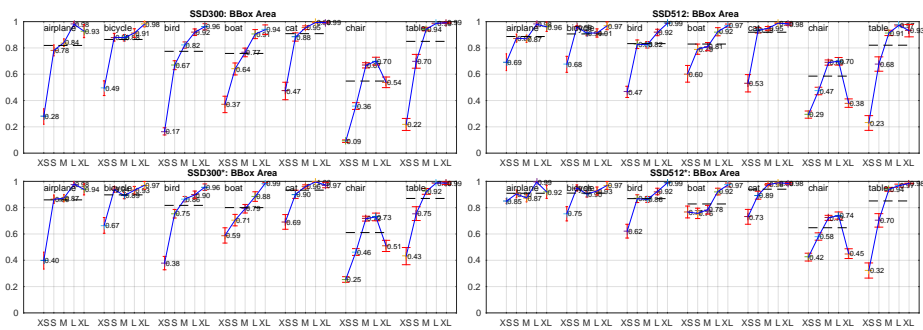
SSD: Single Shot MultiBox Detector   13

Fig. 5: **Detection examples on COCO `test-dev` with SSD512 model.** We show detections with scores higher than 0.6. Each color corresponds to an object category.

| Method | VOC2007 `test` | | VOC2012 `test` | | COCO `test-dev2015` | | |
| | 07+12 | 07+12+COCO | 07++12 | 07++12+COCO | trainval35k | | |
	0.5	0.5	0.5	0.5	0.5:0.95	0.5	0.75
SSD300	74.3	79.6	72.4	77.5	23.2	41.2	23.4
SSD512	76.8	81.6	74.9	80.0	26.8	46.5	27.8
SSD300*	77.2	81.2	75.8	79.3	25.1	43.1	25.8
SSD512*	**79.8**	**83.2**	**78.5**	**82.2**	**28.8**	**48.5**	**30.3**

Table 6: **Results on multiple datasets when we add the image expansion data augmentation trick.** SSD300* and SSD512* are the models that are trained with the new data augmentation.

Fig. 6: **Sensitivity and impact of object size with new data augmentation on VOC2007 `test` set using [21].** The top row shows the effects of BBox Area per category for the original SSD300 and SSD512 model, and the bottom row corresponds to the SSD300* and SSD512* model trained with the new data augmentation trick. It is obvious that the new data augmentation trick helps detecting small objects significantly.

3.7   Inference time

Considering the large number of boxes generated from our method, it is essential to perform non-maximum suppression (nms) efficiently during inference. By using a confidence threshold of 0.01, we can filter out most boxes. We then apply nms with jaccard overlap of 0.45 per class and keep the top 200 detections per image. This step costs about 1.7 msec per image for SSD300 and 20 VOC classes, which is close to the total time (2.4 msec) spent on all newly added layers. We measure the speed with batch size 8 using Titan X and cuDNN v4 with Intel Xeon E5-2667v3@3.20GHz.

Table 7 shows the comparison between SSD, Faster R-CNN[2], and YOLO[5]. Both our SSD300 and SSD512 method outperforms Faster R-CNN in both speed and accuracy. Although Fast YOLO[5] can run at 155 FPS, it has lower accuracy by almost 22% mAP. To the best of our knowledge, SSD300 is the first real-time method to achieve above 70% mAP. Note that about 80% of the forward time is spent on the base network (VGG16 in our case). Therefore, using a faster base network could even further improve the speed, which can possibly make the SSD512 model real-time as well.

4   Related Work

There are two established classes of methods for object detection in images, one based on sliding windows and the other based on region proposal classification. Before the advent of convolutional neural networks, the state of the art for those two approaches – Deformable Part Model (DPM) [26] and Selective Search [1] – had comparable performance. However, after the dramatic improvement brought on by R-CNN [22], which combines selective search region proposals and convolutional network based post-classification, region proposal object detection methods became prevalent.

The original R-CNN approach has been improved in a variety of ways. The first set of approaches improve the quality and speed of post-classification, since it requires

Method	mAP	FPS	batch size	# Boxes	Input resolution
Faster R-CNN (VGG16)	73.2	7	1	~ 6000	$\sim 1000 \times 600$
Fast YOLO	52.7	155	1	98	448×448
YOLO (VGG16)	66.4	21	1	98	448×448
SSD300	74.3	46	1	8732	300×300
SSD512	76.8	19	1	24564	512×512
SSD300	74.3	59	8	8732	300×300
SSD512	76.8	22	8	24564	512×512

Table 7: **Results on Pascal VOC2007 `test`.** SSD300 is the only real-time detection method that can achieve above 70% mAP. By using a larger input image, SSD512 outperforms all methods on accuracy while maintaining a close to real-time speed.

the classification of thousands of image crops, which is expensive and time-consuming. SPPnet [9] speeds up the original R-CNN approach significantly. It introduces a spatial pyramid pooling layer that is more robust to region size and scale and allows the classification layers to reuse features computed over feature maps generated at several image resolutions. Fast R-CNN [6] extends SPPnet so that it can fine-tune all layers end-to-end by minimizing a loss for both confidences and bounding box regression, which was first introduced in MultiBox [7] for learning objectness.

The second set of approaches improve the quality of proposal generation using deep neural networks. In the most recent works like MultiBox [7,8], the Selective Search region proposals, which are based on low-level image features, are replaced by proposals generated directly from a separate deep neural network. This further improves the detection accuracy but results in a somewhat complex setup, requiring the training of two neural networks with a dependency between them. Faster R-CNN [2] replaces selective search proposals by ones learned from a region proposal network (RPN), and introduces a method to integrate the RPN with Fast R-CNN by alternating between fine-tuning shared convolutional layers and prediction layers for these two networks. This way region proposals are used to pool mid-level features and the final classification step is less expensive. Our SSD is very similar to the region proposal network (RPN) in Faster R-CNN in that we also use a fixed set of (default) boxes for prediction, similar to the anchor boxes in the RPN. But instead of using these to pool features and evaluate another classifier, we simultaneously produce a score for each object category in each box. Thus, our approach avoids the complication of merging RPN with Fast R-CNN and is easier to train, faster, and straightforward to integrate in other tasks.

Another set of methods, which are directly related to our approach, skip the proposal step altogether and predict bounding boxes and confidences for multiple categories directly. OverFeat [4], a deep version of the sliding window method, predicts a bounding box directly from each location of the topmost feature map after knowing the confidences of the underlying object categories. YOLO [5] uses the whole topmost feature map to predict both confidences for multiple categories and bounding boxes (which are shared for these categories). Our SSD method falls in this category because we do not have the proposal step but use the default boxes. However, our approach is more flexible than the existing methods because we can use default boxes of different aspect

ratios on each feature location from multiple feature maps at different scales. If we only use one default box per location from the topmost feature map, our SSD would have similar architecture to OverFeat [4]; if we use the whole topmost feature map and add a fully connected layer for predictions instead of our convolutional predictors, and do not explicitly consider multiple aspect ratios, we can approximately reproduce YOLO [5].

5   Conclusions

This paper introduces SSD, a fast single-shot object detector for multiple categories. A key feature of our model is the use of multi-scale convolutional bounding box outputs attached to multiple feature maps at the top of the network. This representation allows us to efficiently model the space of possible box shapes. We experimentally validate that given appropriate training strategies, a larger number of carefully chosen default bounding boxes results in improved performance. We build SSD models with at least an order of magnitude more box predictions sampling location, scale, and aspect ratio, than existing methods [5,7]. We demonstrate that given the same VGG-16 base architecture, SSD compares favorably to its state-of-the-art object detector counterparts in terms of both accuracy and speed. Our SSD512 model significantly outperforms the state-of-the-art Faster R-CNN [2] in terms of accuracy on PASCAL VOC and COCO, while being $3\times$ faster. Our real time SSD300 model runs at 59 FPS, which is faster than the current real time YOLO [5] alternative, while producing markedly superior detection accuracy.

Apart from its standalone utility, we believe that our monolithic and relatively simple SSD model provides a useful building block for larger systems that employ an object detection component. A promising future direction is to explore its use as part of a system using recurrent neural networks to detect and track objects in video simultaneously.

6   Acknowledgment

References

1. Uijlings, J.R., van de Sande, K.E., Gevers, T., Smeulders, A.W.: Selective search for object recognition. IJCV (2013)
2. Ren, S., He, K., Girshick, R., Sun, J.: Faster R-CNN: Towards real-time object detection with region proposal networks. In: NIPS. (2015)
3. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: CVPR. (2016)
4. Sermanet, P., Eigen, D., Zhang, X., Mathieu, M., Fergus, R., LeCun, Y.: Overfeat: Integrated recognition, localization and detection using convolutional networks. In: ICLR. (2014)

5. Redmon, J., Divvala, S., Girshick, R., Farhadi, A.: You only look once: Unified, real-time object detection. In: CVPR. (2016)

6. Girshick, R.: Fast R-CNN. In: ICCV. (2015)

7. Erhan, D., Szegedy, C., Toshev, A., Anguelov, D.: Scalable object detection using deep neural networks. In: CVPR. (2014)

8. Szegedy, C., Reed, S., Erhan, D., Anguelov, D.: Scalable, high-quality object detection. arXiv preprint arXiv:1412.1441 v3 (2015)

9. He, K., Zhang, X., Ren, S., Sun, J.: Spatial pyramid pooling in deep convolutional networks for visual recognition. In: ECCV. (2014)

10. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: CVPR. (2015)

11. Hariharan, B., Arbeláez, P., Girshick, R., Malik, J.: Hypercolumns for object segmentation and fine-grained localization. In: CVPR. (2015)

12. Liu, W., Rabinovich, A., Berg, A.C.: ParseNet: Looking wider to see better. In: ILCR. (2016)

13. Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., Torralba, A.: Object detectors emerge in deep scene cnns. In: ICLR. (2015)

14. Howard, A.G.: Some improvements on deep convolutional neural network based image classification. arXiv preprint arXiv:1312.5402 (2013)

15. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. In: NIPS. (2015)

16. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C., Fei-Fei, L.: Imagenet large scale visual recognition challenge. IJCV (2015)

17. Chen, L.C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L.: Semantic image segmentation with deep convolutional nets and fully connected crfs. In: ICLR. (2015)

18. Holschneider, M., Kronland-Martinet, R., Morlet, J., Tchamitchian, P.: A real-time algorithm for signal analysis with the help of the wavelet transform. In: Wavelets. Springer (1990) 286–297

19. Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T.: Caffe: Convolutional architecture for fast feature embedding. In: MM. (2014)

20. Glorot, X., Bengio, Y.: Understanding the difficulty of training deep feedforward neural networks. In: AISTATS. (2010)

21. Hoiem, D., Chodpathumwan, Y., Dai, Q.: Diagnosing error in object detectors. In: ECCV 2012. (2012)

22. Girshick, R., Donahue, J., Darrell, T., Malik, J.: Rich feature hierarchies for accurate object detection and semantic segmentation. In: CVPR. (2014)

23. Zhang, L., Lin, L., Liang, X., He, K.: Is faster r-cnn doing well for pedestrian detection. In: ECCV. (2016)

24. Bell, S., Zitnick, C.L., Bala, K., Girshick, R.: Inside-outside net: Detecting objects in context with skip pooling and recurrent neural networks. In: CVPR. (2016)

25. COCO: Common Objects in Context. http://mscoco.org/dataset/#detections-leaderboard (2016) [Online; accessed 25-July-2016].

26. Felzenszwalb, P., McAllester, D., Ramanan, D.: A discriminatively trained, multiscale, deformable part model. In: CVPR. (2008)

EXHIBIT A-8

Classification of Quantitative Light-Induced Fluorescence Images Using Convolutional Neural Network

Sultan Imangaliyev[1,3,6] *, Monique H. van der Veen[3], Catherine M. C. Volgenant[3], Bruno G. Loos[3], Bart J. F. Keijser[3,4], Wim Crielaard[3], and Evgeni Levin[5,6]

[1] VU University Medical Center Amsterdam, Amsterdam, The Netherlands.
[2] Cancer Center Amsterdam, Amsterdam, The Netherlands.
[3] Academic Centre for Dentistry Amsterdam, Amsterdam, The Netherlands.
[4] Netherlands Organisation for Applied Scientific Research, Zeist, The Netherlands.
[5] Academic Medical Center, Amsterdam, The Netherlands
[6] Horaizon BV, Rotterdam, The Netherlands
s.imangaliyev@vumc.nl

Abstract. Images are an important data source for diagnosis and treatment of oral diseases. The manual classification of images may lead to misdiagnosis or mistreatment due to subjective errors. In this paper an image classification model based on Convolutional Neural Network is applied to Quantitative Light-induced Fluorescence images. The deep neural network outperforms other state of the art shallow classification models in predicting labels derived from three different dental plaque assessment scores. The model directly benefits from multi-channel representation of the images resulting in improved performance when, besides the Red colour channel, additional Green and Blue colour channels are used.

Keywords: Deep Learning, Convolutional Neural Networks, Bioinformatics, Quantitative Light-Induced Fluorescence

1   Introduction

Diagnosis and therapy in many areas of medicine, including dentistry, nowadays extensively rely on technological advances in biomedical imaging. One of the challenges in the diagnosis of dental patients during daily practice is assessment of their dental plaque level. A novel way to look at this plaque is the use of a Quantitative Light-induced Fluorescence (QLF) camera. When the QLF-camera is used some dental plaque fluoresces red, which is suggested to be an indication for the pathogenicity of the dental plaque [18].

In this paper we apply deep artificial neural network on QLF-images to make a predictive classification model, where class separation is based on the amount

---

* Corresponding author.

2   Imangaliyev et al.

of red fluorescent dental plaque disclosed in such images. Although both intra-examiner and inter-examiner reliability of manual assessment of QLF-images are shown to be high [19], this may become expensive and laborious if the number of images is large. Therefore, there is a need to automate this procedure by implementing a computer-based system for assessment of QLF-images. Existing computer programs developed for this goal have several drawbacks which limit efficiency of QLF-images assessment. They require that the images must have been captured under the fixed circumstances such as camera geometry, focal distance and ambient light conditions [9], which is hard to achieve under clinical settings.

The problem mentioned above could be solved by the use of Deep Learning models, because descriptive features can be learnt directly from raw data representations [10] being insensitive to ambient conditions and natural image variability. Since images have a special two-dimensional structure, a group of Deep Learning methods called Convolutional Neural Network (CNN) explicitly uses the advantages of such a representation [7,12]. Applications of CNN may include both non-biological [2] and biological images [3].

The aim of this paper is to describe the novel application of CNN to QLF-images obtained during clinical intervention study [18]. Furthermore, we compare the performances of the CNN and several state of the art classification models. We tested all of these models on three existing plaque assessment scoring systems. We also checked the influence of adding various colour channels on the model performance. Possible differences were explained based on the biological nature of the problem and based on the properties of these models. Previous studies on this topic either focused on only a single plaque scoring system without providing detailed analysis of results [6] or used small dataset of different images and different network architecture [8].

2   Materials and Methods

2.1   Convolutional Neural Networks

Many of the modern deep learning models utilize very deep architectures to achieve superhuman performance in solving object recognition problems [15,17]. One of such architectures is a novel ultra-deep residual learning network (ResNet) [4]. This architecture can be implemented by adding so called 'shortcut connections' [5] which skip one or more layers. They perform a mapping so that their outputs are added to the outputs of the stacked layes. The whole network can be trained and implemented by using common libraries without modifying the solvers, hence adding neither extra parameters nor computational complexity. ResNet and many other architectures [7,12] use convolutional operator in extracting useful feature mappings in image classification task. Generally, given the filter $K \subset \mathbb{R}^{(2k_1+1) \times (2k_2+1)}$, the discrete convolution of the image I with

filter K is given by

$$(I * K)_{r,s} := \sum_{u=-h_1}^{h_1} \sum_{v=-h_2}^{h_2} K_{u,v} I_{r+u,s+v}. \qquad (1)$$

Let layer $l \in \mathbb{Z}$ be a convolutional layer. The ith feature map in layer l, denoted $Y_i^{(l)}$, is computed as

$$Y_i^{(l)} = B_i^{(l)} + \sum_{j=1}^{m_1^{(l-1)}} K_{i,j}^{(l)} * Y_j^{(l-1)}, \qquad (2)$$

where $B_i^{(l)}$ is a bias matrix and $K_{i,j}^{(l)}$ is the filter of size $(2h_1^{(l)} + 1) \times (2h_2^{(l)} + 1)$ connecting the jth feature map in layer $(l-1)$ with the ith feature map in layer l [12].

2.2   Dataset

The analyzed 427 QLF-images were taken during a clinical intervention study [18] which was conducted at the Academic Centre for Dentistry Amsterdam. Those images were translated into a combined dataset of three colour channels with 216×324 raw pixel intensity values in each of them. In total, three different experiments were performed on labels derived from plaque scoring systems such as Red Fluorescent Plaque Percentage (RF-PP) [18], Red Fluorescent modified Quigley-Hein index (RF-mQH) [19] and modified Silness-Loe Plaque index (mSLP) [20].

2.3   Experimental Setup

The CNN model was implemented on an NVIDIA GeForce GTX Titan X Graphics Processing Unit (GPU) using the *Theano* package [1]. To compare the influence of different colour channels three dataset compositions were tested which are only Red, Red with Green, or full RGB representations. To compare the CNN performance with the performance of the other models, experiments were performed using various shallow classification models implemented in the *Scikit-learn* package [14] such as Logistic Regression (LR), Support Vector Machines Classifier with Gaussian Kernel (SVMC-K), Support Vector Machines Classifier with Linear Kernel (SVMC-L), Gaussian Nave Bayes Classifier (GNB), Gradient Boosting Classifier (GBC), K-Neighbors Classifier (KNC), and Random Forest Classifier (RFC).

Hyperparameters of those models were selected via an exhaustive grid search with stratified shuffled cross-validation procedure so that 80% of the dataset was used as a training set, 10% as a validation set, and the rest 10% as a test set. All binary models were adapted to a multiclass setting by using a *one-versus-all* approach. The predictive performance of the models was assessed by calculating the F_1-score [16]. The reported final F_1-score was obtained by averaging the results of ten random shuffles with fixed test-train splits across all models.

4   Imangaliyev et al.

3   Results and Discussion

3.1   Model Performance Evaluation

Results of experiments for RF-PP, RF-mQH and mSLP labels are provided in Figure 1, Figure 2, and Figure 3 respectively. As it is seen from Figure 1, in the experiment with the RF-PP label, most of the models have a perfect classification performance on the training dataset, but a poor performance on the test dataset. Moreover, the results indicate that using only the Red channel results in a relatively good and comparable performance between both SVM models and Logistic Regression. Adding the Green and especially Blue channels improves the performance of CNN compared to the other models. As a result, the best model (CNN) provided a 0.76 ± 0.05 F_1-score on the test set and a 0.89 ± 0.11 F_1-score on the training set.

Similar to the experiment with RF-PP labels, results depicted in Figure 2, and Figure 3 clearly demonstrate the advantage of CNN over the other models, especially after adding the Green channel. As a result, the best model (CNN) provided a 0.54 ± 0.07 F_1-score on the test set for RF-mQH labels and a 0.40 ± 0.08 F_1-score on the test set for mSLP labels. However, unlike in the RF-PP case, adding the Blue channel did not improve and even decreased the performance for most of the models. Also, there is a clear difference between the performance of models applied on RF-PP and the other labels overall. Namely, even the best model's F_1-scores are in the interval $[0.4, 0.55]$ in the experiments with RF-mQH and mSLP labels, which are much less than the 0.76 achieved in experiments with the RF-PP label.
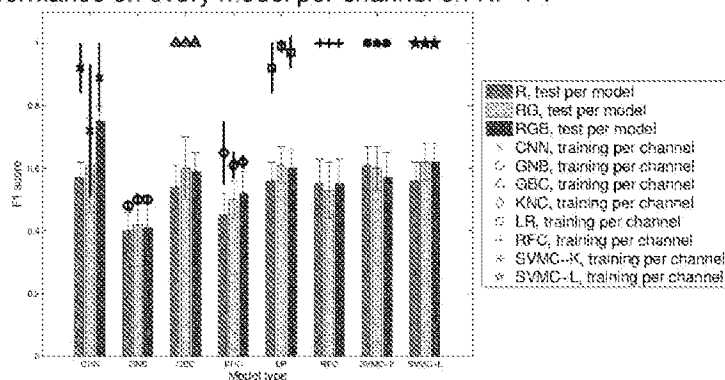
Fig. 1. Test and training performance of models on QLF-images using classes derived from the Red Fluorescent Plaque Percentage (RF-PP) values as labels.

Classification of QLF-images Using Convolutional Neural Network   5
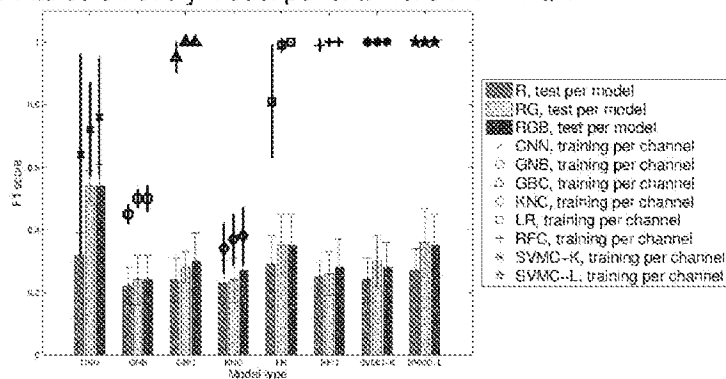
Fig. 2. Test and training performance of models on QLF-images using classes derived from the average Red Fluorescent modified Quigley-Hein (RF-mQH) values as labels.
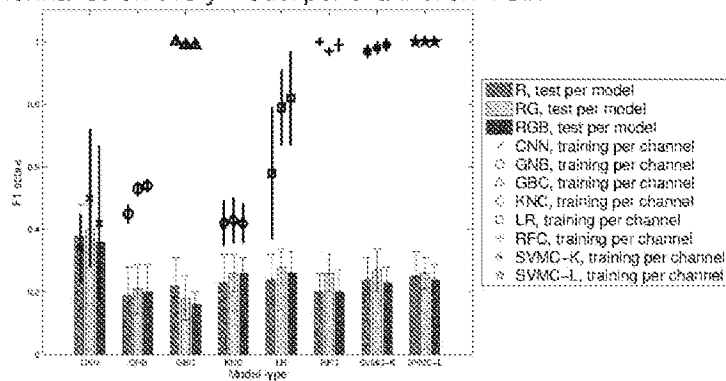
Fig. 3. Test and training performance of models on QLF-images using classes derived from the percentage of modified Silness-Loe Plaque (mSLP) index values as labels.

6   Imangaliyev et al.

3.2   Advantages of the Deep Learning Model

The results of the models' predictive performance evaluation clearly demonstrated advantage of the CNN model over the other models. In general, the predictive performance of the model on previously unseen data, i.e., its generalization can be improved if certain *a priori* information about the problem is added into the choice of the model architecture [11]. In case of images, domain information about the problem can be utilized by a model if such a model is able to learn spatial information between the pixels of an image. This property is explicitly embedded into the CNN model via a discrete convolution operation [10]. In the case of the QLF-images the model may learn, for example, the intensity of red colour associated with plaque, or the sharpness of edges between gingiva and teeth as well as between teeth. Classification results shown in Figures 1, 2, 3 indicate the robustness of CNN to overfitting despite image variability. Other models used in this study do not directly embed spatial information unique for image pixel representation, thus these models have poorer generalization properties and result in a lower classification performance on previously unseen data.

The QLF-images are a good example of images where learning invariant representation is crucial for good predictive performance. Typical examples of QLF-images for each of the three RF-PP classes are provided in Figure 4. As
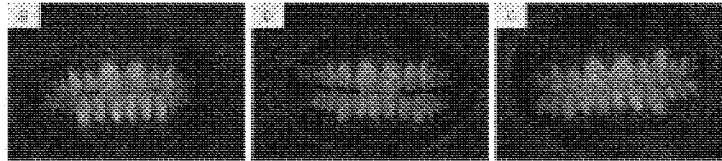
Fig. 4. Typical examples of QLF-images taken at the last day of the clinical intervention separated into three classes depending on different levels of plaque accumulation, for a subject with low (a), moderate (b) or high (c) red fluorescent plaque accumulation.

seen from this figure, these images were taken under various conditions such as slightly different focal distances, rotations, angles and not all images are perfectly centered or focussed to get better resolution. Besides ambient conditions during taking the pictures, the definition of every person is unique. Thus, there is a risk that standard models would overfit and learn variations in angles and distances which are not important for the plaque assessment.

3.3   Influence of Multi-channel Representation

For the experiments on the RF-PP plaque labels, the CNN model results in superior performance over the other classification models if all three colour channels

were used. In the experiments on the RF-mQH and mSLP labels, an improvement was achieved when only the Green channel was added. Moreover, the standard deviation of the training performance tends to be narrower compared to when the model is applied on the Red channel only. This is especially true for GBC, LR and both of the SVMC models.

The Red over Green ratio of pixel values is generally used to identify red fluorescent plaque. Therefore, previous work performed on QLF-images [13,9] used the Red over Green pixel intensities' ratio instead of using the Red channel's pixel intensity values only. The Green channel helps to distinguish plaque from gingiva, since they have slightly different pixel values in Green channel of RGB representation. As for adding the Blue channel, due to technical implementation of the QLF-camera, the blue backscattered light is expected to produce sharper defined edges in images with little red fluorescent plaque, in comparison to images with a thicker plaque. The CNN model incorporates usage of all three colour channels without calculating ratios, thus numerically it is more stable and preferable. Based on these results we conclude that the CNN model benefits from multi-channel representation of the images. Precisely speaking, the CNN model efficiently and explicitly uses the fact that each colour channel contains important information relevant to the classification task.

4   Conclusion

In this study, we applied the CNN model for the automatic classification of red fluorescent dental plaque images. A comparison with several other state of the art shallow classification methods clearly showed the advantage of the CNN model in achieving a higher prediction performance. Such a result was possible because the CNN model directly learns invariant feature representations from raw pixel intensity values without engineering of hand-crafted features. We expect that Deep Learning of red fluorescent dental plaque images can help dental practitioners to perform efficient fluorescent plaque assessments and thus contribute to the improvement of patients' oral health.

References

1. Bergstra, J., Bastien, F., Breuleux, O., Lamblin, P., Pascanu, R., Delalleau, O., Desjardins, G., Warde-Farley, D., Goodfellow, I., Bergeron, A., et al.: Theano: Deep learning on GPUs with Python. In: NIPS 2011, BigLearning Workshop, Granada, Spain (2011)
2. David, O.E., Netanyahu, N.S.: DeepPainter: Painter classification using deep convolutional autoencoders. In: International Conference on Artificial Neural Networks. pp. 20–28. Springer (2016)
3. Esteva, A., Kuprel, B., Novoa, R.A., Ko, J., Swetter, S.M., Blau, H.M., Thrun, S.: Dermatologist-level classification of skin cancer with deep neural networks. Nature 542(7639), 115–118 (2017)

8   Imangaliyev et al.

4. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 770–778 (2016)
5. He, K., Zhang, X., Ren, S., Sun, J.: Identity mappings in deep residual networks. In: European Conference on Computer Vision. pp. 630–645. Springer (2016)
6. Imangaliyev, S., van der Veen, M.H., Volgenant, C.M., Keijser, B.J., Crielaard, W., Levin, E.: Deep learning for classification of dental plaque images. In: International Workshop on Machine Learning, Optimization and Big Data. pp. 407–410. Springer (2016)
7. Jarrett, K., Kavukcuoglu, K., Ranzato, M., LeCun, Y.: What is the best multi-stage architecture for object recognition? In: Computer Vision, 2009 IEEE 12th International Conference on. pp. 2146–2153. IEEE (2009)
8. Kang, J., Li, X., Luan, Q., Liu, J., Min, L.: Dental plaque quantification using cellular neural network-based image segmentation. In: Intelligent computing in signal processing and pattern recognition. pp. 797–802. Springer (2006)
9. Kim, Y.S., Lee, E.S., Kwon, H.K., Kim, B.I.: Monitoring the maturation process of a dental microcosm biofilm using the Quantitative Light-induced Fluorescence-digital (QLF-D). Journal of dentistry 42(6), 691–696 (2014)
10. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature 521(7553), 436–444 (2015)
11. LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., Jackel, L.D.: Backpropagation applied to handwritten zip code recognition. Neural computation 1(4), 541–551 (1989)
12. LeCun, Y., Kavukcuoglu, K., Farabet, C., et al.: Convolutional networks and applications in vision. In: ISCAS. pp. 253–256 (2010)
13. Lee, E.S., Kang, S.M., Ko, H.Y., Kwon, H.K., Kim, B.I.: Association between the cariogenicity of a dental microcosm biofilm and its red fluorescence detected by Quantitative Light-induced Fluorescence-Digital (QLF-D). Journal of dentistry 41(12), 1264–1270 (2013)
14. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al.: Scikit-learn: Machine learning in Python. Journal of Machine Learning Research 12, 2825–2830 (2011)
15. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)
16. Sokolova, M., Lapalme, G.: A systematic analysis of performance measures for classification tasks. Information Processing & Management 45(4), 427–437 (2009)
17. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 1–9 (2015)
18. van der Veen, M.H., Volgenant, C.M., Keijser, B.J., ten Cate, J.B., Crielaard, W.: Dynamics of red fluorescent dental plaque during experimental gingivitis - a cohort study. Journal of dentistry 48, 71–76 (2016)
19. Volgenant, C.M., y Mostajo, M.F., Rosema, N.A., van der Weijden, F.A., ten Cate, J.B., van der Veen, M.H.: Comparison of red autofluorescing plaque and disclosed plaque - a cross-sectional study. Clinical oral investigations 20(9), 2551–2558 (2016)
20. Weijden, G., Timmerman, M., Nijboer, A., Lie, M., Velden, U.: A comparative study of electric toothbrushes for the effectiveness of plaque removal in relation to toothbrushing duration. Journal of clinical periodontology 20(7), 476–481 (1993)

EXHIBIT A-9

Teeth/Palate and Interdental Segmentation Using Artificial Neural Networks

Kelwin Fernandez and Carolina Chang

Grupo de Inteligencia Artificial
Universidad Simón Bolívar
Caracas, Venezuela
kelwin@gia.usb, cchang@usb.ve

Abstract. We present a computational system that combines Artificial Neural Networks and other image processing techniques to achieve teeth/palate segmentation and interdental segmentation in palatal view photographs of the upper jaw. We segment the images into teeth and non-teeth regions. We find the palatal arch by adjusting a curve to the teeth region, and further segment teeth from each other. Best results to detect and segment teeth were obtained with Multilayer Perceptrons trained with the error backpropagation algorithm in comparison to Support Vector Machines. Neural Networks reached up to 87.52% accuracy at the palate segmentation task, and 88.82% at the interdental segmentation task. This is an important initial step towards low-cost, automatic identification of infecto-contagious oral diseases that are related to HIV and AIDS.

Keywords: teeth/palate segmentation, multilayer perceptron, support vector machines.

1   Introduction

Thirty-four million people were living with HIV by the end of 2010 according to the World Health Organization [1]. A 93.5% of these people resided in developing countries, and only 47% of elegible patients in this subgroup received the antiretroviral treatment they needed.

Evidence suggests that about 70% of people with HIV have oral diseases, including Leucoplasia Vellosa, Kaposi's Sarcoma and Candidiasis [2]. The diagnosis of these diseases is very important because some of them may indicate the evolution of HIV towards AIDS [3]. Sadly, the number of dental care centers that treat infecto-contagious deceases in developing countries is limited, and in many cases, insufficient. We believe that some oral diseases that indicate the presence of HIV/AIDS can be diagnosed automatically. A computational, low-cost tool for detecting oral infecto-contagious diseases can lead to health care improvement, especially in low- and middle-income countries.

We focus on the problem of teeth/palate segmentation in palatal view photographs of the upper jaw (maxilla) as an initial step towards diagnosing oral

176   K. Fernandez and C. Chang
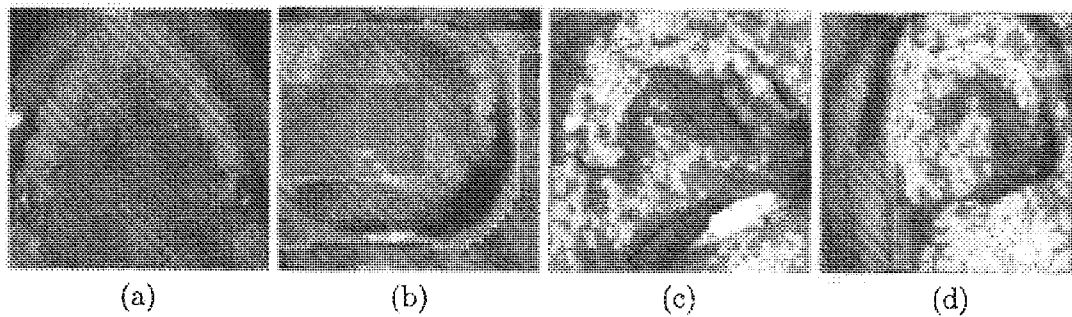
Fig. 1. (a) Healthy Palate. Oral diseases: (b) Candidiasis, (c) Hepatitis, (d) Kaposi's Sarcoma.

diseases. The palate is an important and large area of the mouth where several diseases can be observed. For example, figure 1 shows cases of patients with oral deseases such as Candidiasis, Hepatitis and Kaposi's Sarcoma. In addition to segmenting the palatal region, we aim to segment each tooth of the upper jaw to give a better description of the teeth region. This information could help refine the palatal region, or detect teeth diseases at future versions of the system.

2   Related Work

Methods for interdental segmentation in dental radiographs for general post-mortem identification are [4] [5]. In [4] a neural network based method is proposed for a postmortem identification system by matching image features extracted from dental radiographs. This system tries to match post-mortem and ante-mortem radiographs of a person. It proposes the use of learnable inherent dental image features for tooth-to-tooth image comparisons. In [5] a dental classification and numbering system to segment, classify, and number teeth in dental bitewing radiographs is proposed. Radiographs are enhanced to isolate teeth to regions of interest using image filters. Once teeth are isolated, a support vector machine classifies each tooth to molar or premolar.

In [13] the problem of lip segmentation in color space is handled using color photographies. The proposed solution set a Gaussian model using the hue and saturation value of each pixel within the lip segment. Then, the memberships of lip and non-lip regions are calculated, and the desirable lip region is obtained based on the memberships.

Radiographs have the advantage of having been taken under common standards, with low-variability machines, but provide less information about diseases compared to palatal view photographs. On the other hand, photographs are highly sensitive to phenomena such as illumination, quality of the image, angle of view, and lack of standards, among other. We address these issues on a variety of palatal view photographs, such as those shown in figure 1.

3   System Overview

The input of the system are palatal view photographs of the upper jaw (maxilla) such as those of figure 1. The system is able to determine if an image holds this constraint. Three stages of the system are described: teeth regions detection, palate segmentation, and interdental segmentation.

The teeth regions detection stage classifies each pixel of the image into teeth/non-teeth regions based mainly on color features (see section 4). Some pre and post-processing are required to enhance the image quality and the output. We trained Support Vector Machines and Artificial Neural Networks. Experiments show that Neural Networks reached 85.05% accuracy.

At the palate segmentation stage a curve is adjusted to the palatal arch by means of an iterative heuristic process. Outsider teeth regions are discarded after the palatal arch is found. Not only the shape but also the width of the curve is adjusted to cover teeth and leave noise outside. The palate region is considered to be the area inside the closed curve. We achieved 87.52% accuracy at the palate classification stage (see section 5).

At the final stage, Support Vector Machines as well as Artificial Neural Networks were trained to detect the starting and ending points of each tooth. To do so, the curved teeth region obtained in the previous step is flattened, and then inspected through a sliding window. This process is explained in section 6. Results show that Artificial Neural Networks reached 88.82% accuracy at this task.

For the sake of clarity, experiments and results for each stage of the system are presented at the end of its corresponding section. All experiments were carried out on an Intel Core2Duo 2.4 GHz processor, with 4GB of RAM, running Ubuntu Linux 11.04.

4   Teeth Detection

Color is determinant in our teeth segmentation method. A color enhancement filter improves color differentiation and normalizes the color range of images. Frequently, color images are in RGB format. When operators such as histogram equalization and contrast adjustment are applied over each component separately (red, green and blue), new undesirable colors can show up [5]. Hence, we transform original RGB images to HSV format (hue, saturation, value).

Xiao&Ohya proposed a contrast enhancement filter for color images [6]. They handle the problem of invalid new colors in the resulting image by keeping the Hue component unchanged, where the color information lays. The contrast in the V component (luminance value) is enhanced and finally, the saturation component is improved by histogram equalization.

We propose a filter based on this method. Our variation has the same approach in the Hue and Value components. However, in our variation, the Saturation component distribution is shifted such that the mean saturation is the same in all images. This variation tries to keep the difference in the saturation of each pixel while brings a saturation value suitable for distinguishing colors.

178   K. Fernandez and C. Chang
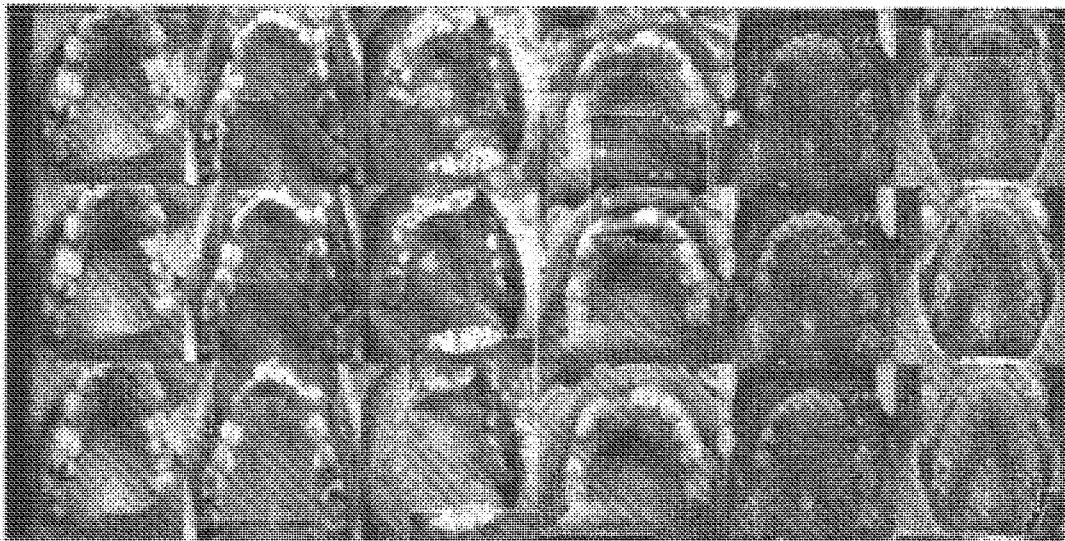
Fig. 2. Image preprocessing. Top: Original images. Middle: Color Enhanced images with Xiao&Ohya's filter. Bottom: Color Enhanced images with the proposed filter.

Finally, in both filters, images are restored to RGB format. These filters are useful in opaque images, where colors are almost indistinguishable. Figure 2 shows the result of applying the filters over six images.

We trained Support Vector Machines and Multilayer perceptrons to classify pixels of the images into teeth/non-teeth regions. SVM kernels include linear, polynomial and RBF kernels. Each SVM was trained varying the parameters using a logarithmic traversing over the search space. The training algorithm of the Multilayer Perceptrons is the classical random sequential back-propagation with a standard sigmoid as the activation function [11]. Neural Networks topologies had one hidden layer, which had from 5 up to 100 hidden neurons.

Experimental results are shown in section 4.1

A post processing filter is applied to erase small blobs of the output image [10]. A blob is defined as a contiguous set of pixels of similar color. Small blobs are treated as noise. Figure 3 shows some examples of the neural network output, and the postprocessing filter output.

4.1   Teeth Detection Experimental Results

A set of 100 images was used. The training and cross-validation set consisted of 10 images, and the test set of 90. Each image had a resolution of 200x200pixels, which means that the training set consisted of 400,000 pixels, and the test set of 3,600,000 pixels.

We wanted to know which image format was best suited for our segmentation task. Therefore, we trained the SVMs and ANNs using each of those formats. RGB, HSV and GRAY refer respectively to the RGB, HSV and gray-scale formats. eRGB and eHSV are the analogous format but with the Xiao&Ohya's enhancement filter, while eRGB', and eHSV' refer to our proposed filter. eGRAY

Teeth/Palate and Interdental Segmentation Using Artificial Neural Networks   179
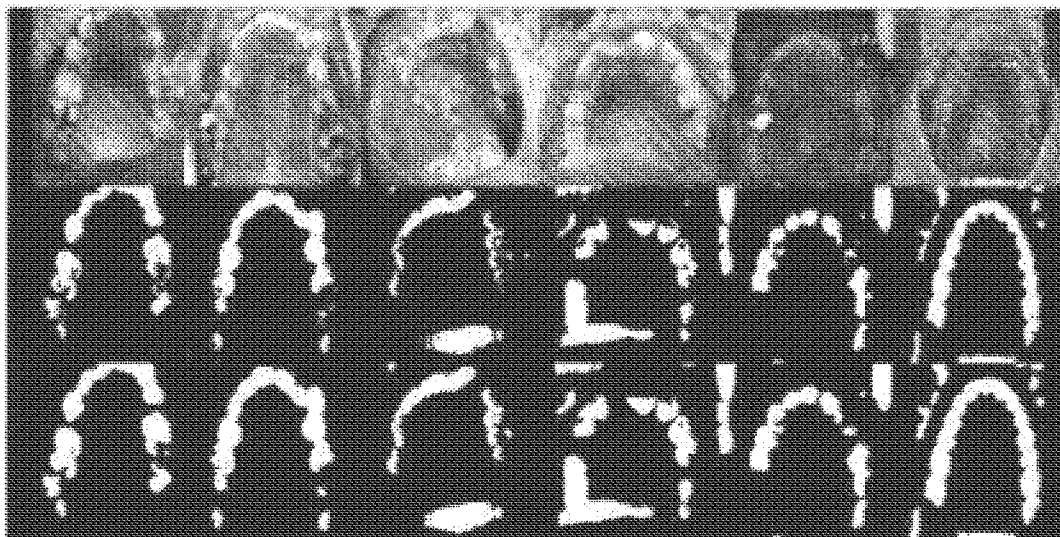
Fig. 3. Teeth Detection. Top: Original Images. Middle: Neural Network output. Bottom: Small blobs filtering.

Table 1. Pixel classification with SVM and ANN

		SVM			ANN			
Argument	Kernel	Train	Test	Time	Top	Train	Test	Time
RGB	RBF	45.34	40.40	0.63	50	82.59	80.02	0.10
eRGB	Linear				100			
eRGB	Linear	56.59	49.69	0.83	5	83.07	78.47	0.08
HSV	Poly					82.78		
eHSV'	Poly	73.00	55.81	0.27	30	85.85	80.67	0.10
eHSV'	RBF							
GRAY	Linear	63.02	72.32	0.57	50	80.88	81.79	0.09
GRAY	Poly	82.10				82.51		

is the gray-scale format with histogram equalization. Each machine receives as input the format of a pixel.

Table 1 shows that Neural Networks work better in this problem than Support Vector Machines. For each color input format, the best Neural Network found outperformed the best Support Vector Machine found. Moreover, the classification time is much smaller for Neural Networks. An interesting result is that the best SVM found receives as argument the enhanced gray-scale image. Unfortunately, this is the most time consuming SVM too.

At the ANN side, it was hard to tell which was the best network found. The best training classification was achieved using the eRGB' input format. However, this network had 100 hidden neurons. Although it was the most time consuming ANN, notice that it is about 7 times faster than the best SVM found. On the other hand, the GRAY format produced the best classification of test images, yet the worst classification over training.

180   K. Fernandez and C. Chang

Table 2. Pixel classification on combined Input Formats for SVM and ANN

		SVM			ANN			
Arguments	Ker_el	Train	Test	Time	hid	Train	Test	Time
eRGB'+eHSV'	Poly	80.89	77.53	0.23	20	**89.02**	81.33	0.10
eRGB'+eHSV'	Linear	71.02	72.34	0.74	5	85.28	82.59	0.09
eRGB'+eGRAY	Linear	61.57	68.87	0.60	5	86.92	79.44	0.09
eRGB'+eGRAY	Poly	80.88	81.79	0.04	30	87.68	80.65	0.10
eHSV'+eGRAY	Poly	75.05	63.30	0.27	25	86.14	80.61	0.09
eHSV'+eGRAY	Linear	71.92	71.41	0.69	40	85.41	81.46	0.10
eRGB'+eHSV'+eGRAY	RBF	65.96	44.93	0.83	25	87.84	84.81	0.11
eRGB'+eHSV'+eGRAY	Linear	80.88	81.79	0.77	40	87.65	85.05	0.11

Table 3. Teeth Classification Rate After Small Blobs Filtering

Stage	Rate (%)	False Positive (%)	False Negative (%)
eRGB' (eHSV') ANN	81.6500	9.8974	8.6726
Blob Filter	86.1185	5.4553	8.4263

To explore further the impact of the image formats, we combined them, as shown in table 2. In general, significant improvements were achieved for SVMs both in classification rate and time as compared with those of table 1. However, none of these SVMs surpassed the 82.19% of classification rate over training achieved by using the eGRAY format only.

The combination of input image formats improved the classification rates of the Neural Networks as well. The best classification rate obtained over training was 89.02%. This result was achieved by a network of 20 hidden neurons, that combined the RGB and HSV images enhanced with our proposed filter. On the other hand, the best classification rate over testing was achieved by a network that combined the RGB, HSV and GRAY images enhanced by the Xiao&Ohya's filter.

From the results we cannot conclude which is the best Neural Network configuration. However, ANNs outperformed SVMs both in classification rate and time.

Table 3 shows the effect of filtering small blobs from the eRGB'+eHSV' Neural Network output. Blob filter leads to an increment of over 4%, decreasing considerably the false positive rate.

5   Palate Segmentation

Once the teeth regions are found it is easier to determine the palate location. We adjust a curve to the teeth regions by means of optimization algorithms [8]. These algorithms perform rotations, translations, scales and deformations of an original curve learned from palate curves.

Teeth/Palate and Interdental Segmentation Using Artificial Neural Networks   181

When skin pixels are classified, some of them have information indistinguishable from teeth, therefore the neural network classifies them as positive examples. In these cases, the curve adjusting algorithm tries to reach a stable point between skin pixels classified as positive and teeth pixels. Cleaning iteratively the noise located outside the curve, corresponding to skin pixels, leads to better results in this stage. This process ends when fixed point is reached by the skin cleaner method.

Fig. 4. Curve adjusted and pixel segmentation after filtering pixels out of the curve. Top: Initial curve. Middle: Curve adjusted to the palatal arch. Bottom: Noise filtered.

As can be seen in figure 4, curves are adjusted within teeth pixels. Picture scale, differences between each individual and noise in the classified image may result in different sizes of teeth. It is needed to detect automatically the bandwidth where teeth are included.

Bandwidth detection should decide for each image the right limits where teeth are covered and noise is outside. An unsupervised and supervised combined method is proposed to solve this problem based in the algorithm exposed in [7].

At the unsupervised step, each boundary pixel classified as positive is assigned to a cluster. Assignment of the point p to a cluster c success if the distance between p and c, $d(p,c)$ is less than certain threshold W and $d(p,c) \leq d(p,c')$ for all c' in the set of clusters, where d is an arbitrary function. In our case d is the squared distance between p and the mass center of c. Otherwise, a new cluster is created with the point p.

When every point is assigned, the algorithm classifies each cluster as normal or anomalous data based on the cluster size. Small clusters contain anomalous data, big clusters contain normal data. The bandwidth selected is the biggest mass center of the normal clusters.

The method described is essentially unsupervised, but there are two variables that should be set, W and normal/anomalous cluster threshold. A supervised method that tries to minimize the distance from actual bandwidth to the selected by the unsupervised algorithm is developed as an initial training method.

When the bandwidth is set, points classified positively between the curve and the allowed bandwidth are denoted as teeth. Points inside the polygon generated by the closed curve are denoted as part of palate, as shown in figure 5

182   K. Fernandez and C. Chang

Fig. 5. Bandwidth selection and teeth/palate segmentation

5.1   Bandwidth Selection Experimental Results

Figure 5 shows the mean relative error of the application of bandwidth selection varying W. For each value of W, normal/anomalous threshold is selected such that minimizes the mean error.

The best results found in the training set are values of 13.3 and 0.1 for W and anomalous threshold. With this selection, the mean relative error between the bandwidth selected and the actual bandwidth is 0.09 in the training set and 0.16 in the test set.

Table 4 shows the teeth classification rate variation after the curve adjustment method and bandwidth selection. Recall from table 3 that the Neural Networks classification rate was 81.33%. Small blob filtering combined with the curve adjustment and bandwidth selection methods, improved the teeth classification rate to 89.74%. There is an important decrease in the percentage of false positives, yet there is an increment in the percentage of false negatives because some teeth regions may not be covered by the palatal curve, as it can be seen in figure 5.

Finally, the classification rate of the pixels as palate is 87.5267%. Errors were mainly false positive (8.7598%). We will work on improving the overall performance of our system.

Table 4. Teeth Classification after Curve Adjustment and Bandwidth Selection

Stage	Rate (%)	False Positive (%)	False Negative (%)
Curve Adjustment	87.8216	3.3103	8.6575
Bandwidth Selection	88.0421	1.8916	10.0663
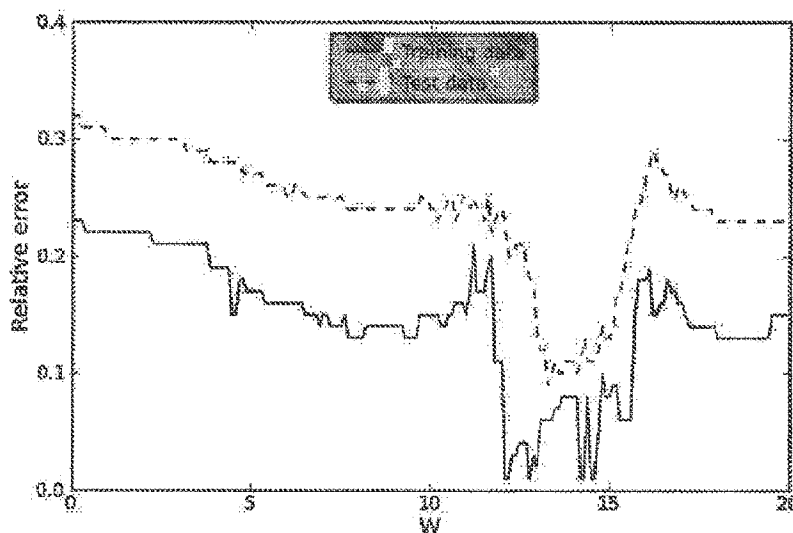
Fig. 6. Bandwidth Selection. Relative error vs. W.

6   Interdental Segmentation

Our first step to segment teeth is to flatten the palatal arch curve, i.e., to transform the teeth band into a rectangle. Unfortunately, some pixels of the resulting rectangle are left blank because of the distortion of the image. Hence, such pixels are filled with a best first algorithm, using as priority the amount of neighboring filled pixels. Each blank pixel is filled with the average color value of its neighbors. Once the rectangle is entirely filled, the image is transformed to gray-scale, and improved by applying histogram equalization [4] [11] and edge enhancement filters [9].

Fig. 7. Interdental Segmentation

184   K. Fernandez and C. Chang

Next, a sliding window algorithm searches for boundary between teeth. The main advantage of having the teeth region represented as a rectangle is that the sliding window can have fixed size. The algorithm complexity is linear in the number of window slides times the complexity time of the classifier.

We tested Support Vector Machine [12] and Multilayer Perceptron trained with Backpropagation [11]. Input are image slices of 50×25 pixels. The classifiers are trained to detect gaps between teeth.

6.1   Interdental Segmentation Experimental Results

ANN outperform SVM at this task as shown in table 5. Best results in time and quality were reached with Neural Networks. The best Neural Network found had 50 neurons at the hidden layer and worked with the original flattened image. A smaller Neural Network of 15 hidden neurons reached the same classification rate during testing using the enhanced images as input. However, this networks did not reach perfect classification rate during training.

Table 5. Interdental Segmentation

		SVM				ANN		
Arguments	Kernel	Train	Test	Time (ms)	Top.	Train	Test	Time (ms)
Original	Linear	86.82	81.14	149	50	100.00	88.82	85
Enhanced	Poly	92.27	83.99	154	15	99.55	88.82	92

7   Conclusions and Future Work

We presented our results on automatic teeth/palate and interdental segmentation in palate view photographs of the upper jaw. Classic Multilayer Perceptrons trained with the error backpropagation algorithm outperformed SVMs for both of these tasks.

Image enhancement, filtering and adaptive curve fitting helped differentiate teeth and palates. Our system was able to cope with variations in patient anatomy, mouth scale in pictures, and view angle. Likewise, it was able to overcome standard problems for image processing systems, such as differences in illumination and image quality.

We wish to make our system more robust, so we are extending its capabilities to detect whether or not the input image depicts the upper jaw. Our final goal is to effectively detect common oral infecto-contagious diseases by looking at the palate, which could help diagnose HIV and AIDS. We believe this is an important initial step in that direction.

References

1. World Health Organization: Progress report 2011: Global HIV/AIDS response Epidemic upd.te and health sector progress towards unive sal access. WHO Press, ISBN 978 92 4 150298 6,
   http://www.who.int/hiv/pub/progress_report2011/en/index.html

2. Barr, C.E.: Dental management of HIV-associated oral mucosal lesions: current and experimental techniques. In: Robertsonn, P.B., Greenspan, J.S. (eds.) Perspectives on Oral Manifestation of AIDS: Diagnosis and Management of HIV-Associated Infections, pp. 77–95. PSG Publishing Co., Inc., Littleton (1988)

3. Arendorf, T.M., Bredekamp, B., Cloete, C., Sauer, G.: Oral manifestation of HIV infection in 600 South African patients. J. Oral Pathol. Med. 27, 176–179 (1998)

4. Nassar, D., Ammar, H.: A neural network system for matching dental radiographs. Pattern Recognition 40, 65–79 (2006)

5. Lin, P., Lai, Y., Huang, P.: An effective classification and numbering system for dental bitewing radiographs using teeth region and contour information. Pattern Recognition (2009)

6. Xiao, D., Ohya, J.: Contrast Enhancement of Color Images Based on Wavelet Transform and Human Visual System. In: Proceedings of the IASTED International Conference Graphics and Visualization in Engineering, Florida (2007)

7. Chimphlee, C., Hanan, A., Noor, M.: Unsupervised Anomaly Detection with Unlabeled Data Using Clustering. In: Proceedings of the Postgraduate Annual Research Seminar (2005)

8. Gendreau, M., Potvin, J.: Handbook of Metaheuristics. Springer (2010)

9. Gonzalez, R., Woods, R.: Digital Image Processing, 3rd edn. Prentice Hall (2007)

10. Shapiro, L., Stockman, G.: Computer Vision. Prentice Hall (2001)

11. LeCun, Y., Bottou, L., Orr, G.B., Müller, K.-R.: Efficient BackProp. In: Orr, G.B., Müller, K.-R. (eds.) Neural Networks: Tricks of the Trade. LNCS, vol. 1524, pp. 9–50. Springer, Heidelberg (1998)

12. Burges, C.: A tutorial on support vector machines for pattern recognition. Knowledge Discovery and Data Mining 2(2) (1998)

13. Li, M., Cheung, Y.-M.: Automatic Segmentation of Color Lip Images Based on Morphological Filter. In: Diamantaras, K., Duch, W., Iliadis, L.S. (eds.) ICANN 2010, Part I. LNCS, vol. 6352, pp. 384–387. Springer, Heidelberg (2010)

EXHIBIT A-10

The AMERICAN HERITAGE® dictionary *of the* English Language

[   ] **Search**

HOW TO USE THE DICTIONARY

To look up an entry in *The American Heritage Dictionary of the English Language,* use the search window above. For best results, after typing in the word, click on the "Search" button instead of using the "enter" key.

Some compound words (like *bus rapid transit, dog whistle,* or *identity theft*) don't appear on the drop-down list when you type them in the search bar. For best results with compound words, place a quotation mark before the compound word in the search window.

guide to the dictionary

THE USAGE PANEL

The Usage Panel is a group of nearly 200 prominent scholars, creative writers, journalists, diplomats, and others in occupations requiring mastery of language. Annual surveys have gauged the acceptability of particular usages and grammatical constructions.

The Panelists

AMERICAN HERITAGE DICTIONARY APP

The new American Heritage Dictionary app is now available for iOS and Android.

THE AMERICAN HERITAGE DICTIONARY BLOG

The articles in our blog examine new words, revised definitions, interesting images from the fifth edition, discussions of usage, and more.

THE 100 WORDS®

See word lists from the best-selling 100 Words Series!

Find out more!

INTERESTED IN DICTIONARIES?

Check out the Dictionary Society of North America at http://www.dictionarysociety.com

vid·e·o 🔊 (vĭd□ē-ō′)

Share: Tweet

n. pl. **vid·e·os**
1.
a. A sequence of images processed electronically into an analog or digital format and displayed on a screen with sufficient rapidity as to create the illusion of motion and continuity.
b. A signal carrying such images.
2.
a. A movie recorded electronically, usually including a soundtrack: *a video of a birthday party.*
b. The presentation of such a work.
c. The electronic medium in which such movies are recorded: *a movie released on video.*
3. A music video.

[From Latin *vidē(re)*, to see; see VIDE + _O_ (on the model of AUDIO).]

vid□e·o *adj.*

Indo-European & Semitic Roots Appendices

Thousands of entries in the dictionary include etymologies that trace their origins back to reconstructed proto-languages. You can obtain more information about these forms in our online appendices:

Indo-European Roots

Semitic Roots

The Indo-European appendix covers nearly half of the Indo-European roots that have left their mark on English words. A more complete treatment of Indo-European roots and the English words derived from them is available in our Dictionary of Indo-European Roots.

American Heritage Dictionary Products

The American Heritage Dictionary, 5th Edition

The American Heritage Dictionary of Idioms

The American Heritage Roget's Thesaurus

Curious George's Dictionary

The American Heritage Children's Dictionary

CONTACT US
- Customer Service
- Make Me An Author
- Ebooks Help with Glose Reader

ABOUT US
- Company Profile
- Leadership Team
- Corporate Social Responsibility
- HarperCollins Careers
- HarperCollins Imprints
- HarperGreen
- Social Media Directory
- Accessibility

FOR READERS
- Browse Reading Guides

FOR AUTHORS
- Submit a Manuscript
- Report Piracy
- Agent Portal

MEDIA
- Publicity Contacts
- Press Room

SERVICES
- HarperCollins Speakers Bureau
- Library Services
- Academic Services
- Desk & Exam Copies
- Review Copies
- OpenBook API
- Marketing Partnerships

COVID-19 RESOURCES & PERMISSIONS
- Permissions for Adult Online Readings
- Permissions for Kids Online Readings

SALES & RIGHTS
- Booksellers & Retailer Ordering
- HarperCollins Catalogs
- Permissions
- Subsidiary Rights
- Media Rights and Content Development

GLOSE APP
- iPhone
- Android

GLOBAL DIVISIONS
- HarperCollins US
- HarperCollins Canada
- HarperCollins Christian
- HarperCollins Australia
- HarperCollins India
- HarperCollins UK

HarperCollins*Publishers*

News Corp

*This website is best viewed in Chrome, Firefox, Microsoft Edge, or Safari. Some characters in pronunciations and etymologies cannot be displayed properly in Internet Explorer.

EXHIBIT A-11

Trending Now   Data Structures   Algorithms   Foundational Courses   Data Science   Practice Problem   Python   Machine Learning   JavaScript   Java   C   C++   ReactJS   NodeJS   Web Development   Web

Difference Between JPEG and MPEG

Read   Discuss   Courses   ⋮

Pre-requisites: JPEG, MPFG

JPEG stands for Joint Photographic Experts Group. .jpg and .jpeg extensions are used to store images in this format. It uses a lossy compression algorithm, which means some of the image data is lost.
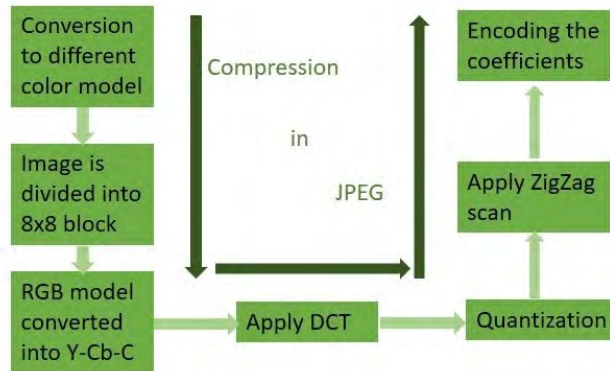
Advantages of JPEG:

- Small file size: JPEG images can be compressed to small file sizes, making them ideal for sharing over the internet or storing on limited storage devices.
- Widely supported: JPEG is a widely supported image format, and most devices and software can read and display JPEG images.
- High-quality output: JPEG compression allows for high-quality output images, even after compression.
- Adjustable compression level: JPEG compression level can be adjusted to balance file size and image quality.

Disadvantages of JPEG:

- Lossy compression: JPEG compression is lossy, which means that some data is lost during the compression process, resulting in a lower quality image.
- Limited editing: Editing a JPEG image repeatedly can result in further loss of quality due to the lossy compression.
- Not ideal for text or graphics: JPEG compression is not well-suited for compressing text, graphics, or images with sharp edges, as these can become blurry or pixelated after compression.

MPEG stands for Moving Picture Experts Group is a standard used for compressing digital video files. It is also a lossy compression method, but it is optimized for compressing moving images rather than still images.

Advantages of MPEG:

- Small file size: MPEG compression can reduce the size of a video file significantly, making it easier to store and share.
- High-quality output: MPEG compression can produce high-quality video output, even after compression.
- Adjustable compression level: MPEG compression level can be adjusted to balance file size and video quality.
- Supports various resolutions: MPEG supports various resolutions, from standard definition to high definition, allowing for flexibility in video production.

Disadvantages of MPEG:

- Lossy compression: MPEG compression is lossy, which means that some data is lost during the compression process, resulting in a lower quality video.
- Limited editing: Editing an MPEG video repeatedly can result in further loss of quality due to the lossy compression.
- Complex compression: MPEG compression is a complex process that requires significant computing power, making it challenging for some devices to handle.

Similarities between JPEG and MPEG :

- Lossy Compression: Both JPEG and MPEG use lossy compression techniques, which means that they discard some information to reduce file size. This results in a reduction in image and video quality compared to the original content.
- Popular Formats: Both JPEG and MPEG are popular formats used for digital content. JPEG is commonly used for still images, while MPEG is used for video and audio.
- Adjustable Compression: Both formats offer adjustable compression levels, allowing users to choose between smaller file sizes or higher image and video quality.
- Widely Supported: Both formats are widely supported by devices and software, making them accessible and easy to use for a variety of purposes.
- Standards-based: Both formats are based on industry standards, which ensures interoperability and compatibility with different devices and software applications.

Difference between JPEG and MPEG

JPEG	MPEG
JPEG stands for Joint Photographic Experts Group.	MPEG stands for Moving Picture Experts Group.
JPEG is primarily used for Web images, digital cameras, other devices that capture still images.	MPEG is primarily used for Movies, TV shows, video clips on the web.
JPEG is mainly used for image compression.	MPEG has various standards for audio and video compression.
Compression ratio of JPEG files is typically 10:1.	Compression ratio of MPEG files can vary from 20:1 or higher depending on the video content.
Compressed as a single JPEG File.	Compressed as a series of frames using inter frame compression.
The extensions used are .jpg and .jpeg.	The extensions used are .mp3 and .mp4.

Whether you're preparing for your first job interview or aiming to upskill in this ever-evolving tech landscape, GeeksforGeeks Courses are your key to success. We provide top-quality content at affordable prices, all geared towards accelerating your growth in a time-bound manner. Join the millions we've already empowered, and we're here to do the same for you. Don't miss out - check it out now!

Last Updated : 05 Apr, 2023   👍 4   🔖

Similar Reads

Difference between JPEG and PNG	Difference between JPEG and GIF
Difference between SVG and JPEG	Difference between JPEG and Bitmap
Difference between JPEG and PSD	Difference between Difference Engine and Analytical Engine
Difference Between Electric Potential and Potential Difference	Difference between Voltage Drop and Potential Difference
Difference and Similarities between PHP and C	Difference between Stop and Wait protocol and Sliding Window protocol

‹ Previous
Difference between Virtualization and Emulation

Next ›
Difference Between Digital Audio and MIDI

Article Contributed By :

mihirpanchal5400
mihirpanchal5400
+ Follow

| Easy | Normal | Medium | Hard | Expert |

Improved By :   _alok_14

Article Tags :   Computer Graphics ,   Difference Between

Improve Article   Report Issue

GeeksforGeeks

A-143, 9th Floor, Sovereign Corporate Tower, Sector-136, Noida, Uttar Pradesh - 201305

feedback@geeksforgeeks.org

Company
About Us
Legal
Terms & Conditions
Careers
In Media
Contact Us
Advertise with us
GFG Corporate Solution
Placement Training Program
Apply for Mentor

Explore
Job-A-Thon Hiring Challenge
Hack-A-Thon
GfG Weekly Contest
Offline Classes (Delhi/NCR)
DSA in JAVA/C++
Master System Design
Master CP
GeeksforGeeks Videos

Languages
Python
Java
C++
PHP
GoLang
SQL
R Language
Android Tutorial

DSA Concepts
Data Structures
Arrays
Strings
Linked List
Algorithms
Searching
Sorting
Mathematical
Dynamic Programming

DSA Roadmaps
DSA for Beginners
Basic DSA Coding Problems
DSA Roadmap by Sandeep Jain
DSA with JavaScript
Top 100 DSA Interview Problems
All Cheat Sheets

Web Development
HTML
CSS
JavaScript
Bootstrap
ReactJS
AngularJS
NodeJS

Lodash

Web Design

Computer Science

GATE CS Notes

Operating Systems

Computer Network

Database Management System

Software Engineering

Digital Logic Design

Engineering Maths

Data Science & ML

Data Science With Python

Data Science For Beginner

Machine Learning Tutorial

Maths For Machine Learning

Pandas Tutorial

NumPy Tutorial

NLP Tutorial

Deep Learning Tutorial

Competitive Programming

Top DSA for CP

Top 50 Tree Problems

Top 50 Graph Problems

Top 50 Array Problems

Top 50 String Problems

Top 50 DP Problems

Top 15 Websites for CP

Interview Corner

Company Wise Preparation

Preparation for SDE

Experienced Interviews

Internship Interviews

Competitive Programming

Aptitude Preparation

Puzzles

Commerce

Accountancy

Business Studies

Economics

Human Resource Management (HRM)

Management

Income Tax

Finance

Statistics for Economics

SSC/ BANKING

SSC CGL Syllabus

SBI PO Syllabus

SBI Clerk Syllabus

IBPS PO Syllabus

IBPS Clerk Syllabus

Aptitude Questions

SSC CGL Practice Papers

Python

Python Programming Examples

Django Tutorial

Python Projects

Python Tkinter

OpenCV Python Tutorial

Python Interview Question

DevOps

Git

AWS

Docker

Kubernetes

Azure

GCP

System Design

What is System Design

Monolithic and Distributed SD

Scalability in SD

Databases in SD

High Level Design or HLD

Low Level Design or LLD

Crack System Design Round

System Design Interview Questions

GfG School

CBSE Notes for Class 8

CBSE Notes for Class 9

CBSE Notes for Class 10

CBSE Notes for Class 11

CBSE Notes for Class 12

English Grammar

UPSC

Polity Notes

Geography Notes

History Notes

Science and Technology Notes

Economics Notes

Important Topics in Ethics

UPSC Previous Year Papers

Write & Earn

Write an Article

Improve an Article

Pick Topics to Write

Share your Experiences

Internships

EXHIBIT A-12

Collins
ENGLISH
DICTIONARY

If you would like to comment on any aspect of this book,
please contact us at the given address or online.
E-mail: dictionaries@harpercollins.co.uk
 facebook.com/collinsdictionary
 @collinsdict

About the type

The A-Z of this dictionary is typeset in CollinsFedra, a
special version of the Fedra family of types designed by
Peter Bil'ak. CollinsFedra has been customized specially
for Collins dictionaries; it includes both sans serif (for
headwords) and serif (entries) versions, in several different
weights. Its large x-height, its open 'eye', and its basis in
the tradition of humanist letterforms make CollinsFedra
both familiar and easy to read at small sizes. It has been
designed to use the minimum space without sacrificing
legibility, as well as including a number of characters and
signs that are specific to dictionary typography. Its
companion phonetic type is the first of its kind to be
drawn according to the same principles as the regular
typeface, rather than assembled from rotated and
reflected characters from other types. Peter Bil'ak
(born 1973, Slovakia) is a graphic and type designer living
in the Netherlands. As well as the Fedra family, he has
designed several other typefaces including Eureka.
His typotheque.com website has become a focal point for
research and debate around contemporary type design.

Bowen ('bəʊən) n Elizabeth (Dorothea Cole). 1899–1973, British novelist and short-story writer, born in Ireland. Her novels include *The Death of the Heart* (1938) and *The Heat of the Day* (1949)

bower¹ ('baʊə) n 1 a shady leafy shelter or recess, as in a wood or garden; arbour 2 *literary* a lady's bedroom or apartments, esp in a medieval castle; boudoir 3 *literary* a country cottage, esp one regarded as charming or picturesque [Old English *būr* dwelling; related to Old Norse *būr* pantry, Old High German *būr* dwelling] ›**bowery** adj ■ **bowerbird** ('baʊə,bɜːd) n 1 any of various songbirds of the family Ptilonorhynchidae, of Australia and New Guinea. The males build bower-like display grounds in the breeding season to attract the females 2 *informal, chiefly Austral* a person who collects miscellaneous objects ■ **bowerwoman** ('baʊə,wʊmən) n, pl -women *archaic* a chamber-woman

bower² ('baʊə) n *nautical* a vessel's bow anchor [C18: from BOW³ + -ER¹]

bower³ ('baʊə) n a jack in euchre and similar card games [C19: from German *Bauer* peasant, jack (in cards)]

Bowery ('baʊərı) n the Bowery a street in New York City noted for its cheap hotels and bars, frequented by vagrants and drunks [C17: from Dutch *bouwerij*, from *bouwen* to farm + *-erij* -ERY; see BOOR, BOER]

Bowie n 1 ('baʊı, 'bəʊı) David, real name David Jones. 1947–2016, British rock singer, songwriter, and film actor. His recordings include "Space Oddity" (1969), *The Rise and Fall of Ziggy Stardust and the Spiders from Mars* (1972), *Heroes* (1977), *Let's Dance* (1983), and *Heathen* (2002) 2 ('bəʊı) James, known as *Jim Bowie*. 1796–1836, US frontiersman. A hero of the Texas Revolution against Mexico (1835–36), he died at the Battle of the Alamo ■ **bowie knife** ('bəʊı) n a stout hunting knife with a short hilt and a guard for the hand [C19: named after Jim BOWIE, who popularized it]

bowl¹ ('bəʊl) n 1 a round container open at the top, used for holding liquid, keeping fruit, serving food, etc 2 Also: **bowlful** the amount a bowl will hold 3 the rounded or hollow part of an object, esp of a spoon or tobacco pipe 4 any container shaped like a bowl, such as a sink or lavatory 5 *chiefly US* a bowl-shaped building or other structure, such as a football stadium or amphitheatre 6 a bowl-shaped depression of the land surface. See also **dust bowl** 7 *literary* a a drinking cup b intoxicating drink [Old English *bolla*; related to Old Norse *bolli*, Old Saxon *bollo*] ■ **bowlful** ('bəʊlfʊl) n amount held by a bowl ■ **bowllike** ('bəʊl,laɪk) adj resembling a bowl; bowl-shaped

bowl² ('bəʊl) n 1 a wooden container used in the game of bowls, having flattened sides, one side usually being flatter than the other in order to make it run on a curved course 2 a large heavy ball with holes for gripping with the fingers and thumb, used in tenpin bowling › vb 3 to roll smoothly or cause to roll smoothly, esp by throwing underarm along the ground 4 (*intr*; usually foll by *along*) to move easily and rapidly, as in a car 5 *cricket* a to send (a ball) down the pitch from one's hand towards the batsman, keeping the arm straight while doing so b Also: **bowl out** to dismiss (a batsman) by delivering a ball that breaks his wicket 6 (*intr*) to play bowls or tenpin bowling 7 (*tr*) (in tenpin bowling) to score (a specified amount): *he bowled 120* [C15: from French *boule*, ultimately from Latin *bulla* bubble] ■ **bowler** ('bəʊlə) n 1 one who bowls in cricket 2 a player at the game of bowls ■ **bowling** ('bəʊlɪŋ) n 1 any of various games in which a heavy ball is rolled down a special alley, usually made of wood, at a group of wooden pins, esp the games of tenpin bowling (tenpins) and skittles (ninepins) 2 the game of bowls 3 *cricket* the act of delivering the ball to the batsman 4 (*modifier*) of or relating to bowls or bowling: *a bowling team* ■ **bowling alley** n 1 a long narrow wooden lane down which the ball is rolled in tenpin bowling b a similar lane or alley, usually with raised sides, for playing skittles (ninepins) 2 a building having several lanes for tenpin bowling ■ **bowling crease** n *cricket* a line marked at the wicket, over which a bowler must not advance fully before delivering the ball ■ **bowling green** n an area of closely mown turf on which the game of bowls is played ■ **bowl over** vb (*tr, adverb*) 1 *informal* to surprise (a person) greatly, esp in a pleasant way; astound; amaze: *he was bowled over by our gift* 2 to knock (a person or thing) down; cause to fall over ■ **bowls** (bəʊlz) n (*functioning as singular*) 1 a a game played on a bowling green in which a small bowl (the jack) is pitched from a mark and two opponents or opposing teams take turns to roll biased wooden bowls towards it, the object being to finish as near the jack as possible b (*as modifier*): *a bowls tournament* 2 skittles or tenpin bowling

bowler¹ ('bəʊlə) or **bowler hat** n a stiff felt hat with a rounded crown and narrow curved brim. US and Canadian name: **derby** [C19: named after John Bowler, 19th-century London hatter]

bowler² ('bəʊlə) n *Dublin dialect* a dog [perhaps from B(OW-WOW) + (H)OWLER]

Bowles (bəʊlz) n Paul. 1910–99, US novelist, short-story writer, and composer, living in Tangiers. His novels include *The Sheltering Sky* (1949) and *The Spider's House* (1955)

bowlingual (,bəʊˈlɪŋgwəl) n a device that allegedly translates a dog's barks and grunts into a human language [C21: from BOW(WOW) + (BI)LINGUAL]

bowr ('bəʊə) n *obsolete* a muscle

bowse (baʊz) vb a variant spelling of **bouse**

bowser ('baʊzə) n 1 a tanker containing fuel for aircraft, military vehicles, etc 2 *Austral and NZ obsolete* a petrol pump [originally a US proprietary name, from S. F. *Bowser*, US inventor, who made the first one in 1885]

bowsie ('baʊzı) or **bowsey** n *Irish informal* 1 a low-class mean or obstreperous person 2 a drunkard [of unknown origin]

Bow Street runner (bəʊ) n (in Britain from 1749 to 1829) an officer at Bow Street magistrates' court, London, whose duty was to pursue and arrest criminals

bow-wow ('baʊ,waʊ, -'waʊ) n 1 a child's word for dog 2 an imitation of the bark of a dog › vb 3 (*intr*) to bark or imitate a dog's bark

bowyangs ('bəʊjæŋz) pl n *Austral and NZ history* a pair of strings or straps secured round each trouser leg below the knee, worn esp by sheep-shearers and other labourers [C19: from English dialect *bowy-yanks* leggings] ■ **bowyang** ('bəʊjæŋ) n one of a pair of bowyangs

box¹ (bɒks) n 1 a a receptacle or container made of wood, cardboard, etc, usually rectangular and having a removable or hinged lid 2 Also called: **boxful** the contents of such a receptacle or the amount it can contain: *he ate a whole box of chocolates* 3 any of various containers for a specific purpose: *a money box; letter box* 4 (*often in combination*) any of various small cubicles, kiosks, or shelters: *a telephone box or callbox; a sentry box; a signal box on a railway* 5 a separate compartment in a public place for a small group of people, as in a theatre or certain restaurants 6 an enclosure within a courtroom. See **jury box**, **witness box** 7 a compartment for a horse in a stable or vehicle. See **loosebox**, **horsebox** 8 *Brit* a small country house occupied by sportsmen when following a field sport, esp shooting 9 a a protective housing for machinery or mechanical parts b the contents of such a box c (*in combination*): *a gearbox* 10 a shaped device of light tough material worn by sportsmen to protect the genitals, esp in cricket 11 a a section of printed matter on a page, enclosed by lines, a border, or white space 12 a central agency to which mail is addressed and from which it is collected or redistributed: *a post-office box; to reply to a box number in a newspaper advertisement* 13 the central part of a computer or the casing enclosing it 14 short for **penalty box** 15 *baseball* either of the designated areas in which the batter may stand 16 the raised seat on which the driver sits in a horse-drawn coach 17 *NZ* a wheeled container for transporting coal in a mine 18 *Austral and NZ* an accidental mixing of herds or flocks 19 a hole cut into the base of a tree to collect the sap 20 short for **Christmas box** 21 a device for dividing water into two or more ditches in an irrigation system 22 an informal name for a **coffin** 23 *taboo, slang* the female genitals 24 **be a box of birds** *NZ* to be very well indeed 25 **the box** *Brit informal* television 26 **think outside the box** or **think out of the box** to think in a different, innovative, or original manner, esp with regard to business practices, products, systems, etc 27 **tick all the boxes** to satisfy all of the apparent requirements for success 28 **out of the box** *Austral informal* outstanding or excellent: *a day out of the box* › vb 29 (*tr*) to put into a box 30 (*tr*; usually foll by *in* or *up*) to prevent from moving freely; confine 31 (*tr*; foll by *in*) *printing* to enclose (text) within a ruled frame 32 (*tr*) to make a cut in the base of (a tree) in order to collect the sap 33 (*tr*) *Austral and NZ* to mix (flocks or herds) accidentally 34 (*tr*; sometimes foll by *up*) *NZ* to confuse: *I am all boxed up* 35 *nautical* short for **boxhaul** 36 **box the compass** *nautical* to name the compass points in order [Old English *box*, from Latin *buxus* from Greek *puxos* BOX³] ›**box,like** adj ■ **boxball** ('bɒks,bɔːl) n US a street game played with a small ball ■ **box beam** n another name for **box girder** ■ **boxboard** ('bɒks,bɔːd) n a tough paperboard made from wood and wastepaper pulp: used for making boxes, etc ■ **box camera** n a simple box-shaped camera having an elementary lens, shutter, and viewfinder ■ **box canyon** n *Western US* a canyon with vertical or almost vertical walls ■ **boxcar** ('bɒks,kɑː) n US and Canadian a closed railway freight van ■ **box chronometer** n *nautical* a ship's chronometer, supported on gimbals in a wooden box ■ **box coat** n 1 a plain short coat that hangs loosely from the shoulders 2 a heavy overcoat, worn formerly by coachmen ■ **box cutter** n a knife-like tool with a short retractable blade ■ **boxed** (bɒkst) adj packaged together for sale in a presentation box ■ **boxed set** n another name for **box set** ■ **box file** n a rigid file which opens like a box, usually made of strong cardboard and able to hold a large quantity of documents ■ **box-fish** ('bɒks,fɪʃ) n, pl -**fish** or -**fishes** another name for **trunkfish** ■ **box-fresh** adj unused or unspoiled; straight from the packaging ■ **boxful** ('bɒksfʊl) n the contents of a box or the amount a box can contain ■ **box girder** n a a girder that is hollow and square or rectangular in shape b (*as modifier*): *a box-girder bridge*. Also called: **box beam** or **boxhaul** ('bɒks,hɔːl) vb *nautical* to bring (a square-rigger) onto a new tack by backwinding the foresails and steering hard round ■ **boxily** ('bɒksɪlı) adv in a boxy manner ■ **boxiness** ('bɒksınəs) n the quality of being boxy ■ **box jellyfish** n any of various highly venomous jellyfishes of the order Cubomedusae, esp *Chironex fleckeri*, of Australian tropical waters, having a cuboidal body with tentacles hanging from each of the lower corners. Also called (Austral): **sea wasp** ■ **box junction** n (in Britain) a road junction having yellow cross-hatching painted on the road surface. Vehicles may only enter the hatched area when their exit is clear ■ **boxkeeper** ('bɒks,kiːpə) n an attendant responsible for theatre boxes ■ **box kite** n a kite with a boxlike frame open at both ends ■ **box number** n 1 the number of an individual pigeonhole at a newspaper to which replies to an advertisement may be addressed 2 the number of an individual pigeonhole at a post office from which mail may be collected ■ **box office** n 1 an office at a theatre, cinema, etc, where tickets are sold 2 the receipts from a play, film, etc 3 a the public appeal of an actor or production: *the musical was bad box office* b (*as modifier*): *a box-office success* ■ **box pleat** n a flat double pleat made by folding under the fabric on either side of it ■ **boxplot** ('bɒks,plɒt) n *statistics* a graphical representation of numerical data consisting of a

B

rectangular box with lines extending from each end ■**boxroom** ('bɒks,ru:m, -,rom) *n* a small room or large cupboard in which boxes, cases, etc, may be stored ■**box seat** *n* **1** a seat in a theatre box **2** in the box seat *Brit, Austral and NZ* in the best position ■**box set** *n* a collection of items of the same type, packaged together for sale in a presentation box ■**box spanner** *n* a spanner consisting of a steel cylinder with a hexagonal end that fits over a nut: used esp to turn nuts in positions that are recessed or difficult of access ■**box spring** *n* a coiled spring contained in a boxlike frame, used as a base for mattresses, chairs, etc ■**box-ticking** *n derogatory* the process of satisfying bureaucratic administrative requirements rather than assessing the actual merit of something ■**boxwallah** ('bɒks,wɒlə) *n derogatory* an itinerant pedlar or salesman in India ■**boxy** ('bɒksɪ) *adj* squarish or chunky in style or appearance: *a boxy square-cut jacket*

box² (bɒks) *vb* **1** (*tr*) to fight (an opponent) in a boxing match **2** (*intr*) to engage in boxing **3** (*tr*) to hit (a person) with the fist; punch or cuff **4** box clever to behave in a careful and cunning way ▷ *n* **5** a punch with the fist, esp on the ear [C14: of uncertain origin; perhaps related to Dutch *boken* to shunt, push into position] ■**boxer** ('bɒksə) *n* **1** a person who boxes, either professionally or as a hobby; pugilist **2** a medium-sized smooth-haired breed of dog with a short nose and a docked tail ■**boxercise** ('bɒksə,saɪz) *n* a system of sustained exercises combining boxing movements with aerobic activities ■**boxer shorts** *pl n* men's underpants shaped like shorts but having a front opening. Also called: **boxers** ■**boxing** ('bɒksɪŋ) *n* a the act, art, or profession of fighting with the fists, esp the modern sport practised under Queensberry rules **b** (*as modifier*): *a boxing enthusiast* ■**boxing glove** *n* one of a pair of thickly padded mittens worn for boxing

box³ (bɒks) *n* **1** a dense slow-growing evergreen tree or shrub of the genus *Buxus*, esp *B. sempervirens*, which has small shiny leaves and is used for hedges, borders, and garden mazes: family *Buxaceae* **2** the wood of this tree. See **boxwood** (1) **3** any of several trees the timber or foliage of which resembles this tree, esp various species of *Eucalyptus* with rough bark [Old English, from Latin *buxus*] ■**boxberry** ('bɒksbərɪ) *n, pl* **-ries 1** the fruit of the partridgeberry or wintergreen **2** another name for **partridgeberry**, **wintergreen** (1) ■**box elder** *n* a medium-sized fast-growing widely cultivated North American maple, *Acer negundo*, which has compound leaves with lobed leaflets. Also called: **ash-leaved maple** ■**boxen** ('bɒksən) *adj archaic* of or relating to the box-tree; made of box-wood ■**boxthorn** ('bɒks,θɔːn) *n* another name for **matrimony vine** ■**boxwood** ('bɒks,wʊd) *n* **1** the hard close-grained yellow wood of the box tree, used to make tool handles, small turned or carved articles, etc **2** the box tree

box calf *n* black calfskin leather, tanned with chromium salts, having a pattern of fine creases formed by boarding [C20: named after Joseph Box, London shoemaker]

Boxer ('bɒksə) *n* **a** a member of a nationalistic Chinese secret society that led an unsuccessful rebellion in 1900 against foreign interests in China **b** (*as modifier*): *the Boxer Rebellion* [C18: rough translation of Chinese *I Ho Ch'üan*, literally: virtuous harmonious fist, altered from *I Ho T'uan* virtuous harmonious society]

Boxgrove man ('bɒksɡrəʊv) *n* a type of primitive man, probably *Homo heidelbergensis*, and probably dating from the Middle Palaeolithic period some 500 000 years ago; remains were found at Boxgrove in West Sussex in 1993 and 1995. See also **Heidelberg man**

Boxing Day *n Brit* the first day (traditionally and strictly, the first weekday) after Christmas, observed as a holiday [C19: from the custom of giving Christmas boxes to tradesmen and staff on this day]

boxty ('bɒkstɪ) *n Irish* a potato pancake

boy (bɔɪ) *n* **1** a male child; lad; youth **2** a man regarded as immature or inexperienced: *he's just a boy when it comes to dealing with women* **3** See **old boy 4** *informal* a group of men, esp a group of friends **5** *usually derogatory* (esp in former colonial territories) a Black person or native male servant of any age **6** *Austral* a jockey or apprentice **7** short for **boyfriend 8** boys will be boys youthful indiscretion or exuberance must be expected and tolerated **9** jobs for the boys *informal* appointment of one's supporters to posts, without reference to their qualifications or ability **10** the boy *Irish informal* the right tool for a particular task: *that's the boy to cut it* ▷ *interj* **11** an exclamation of surprise, pleasure, contempt, etc: *boy, is he going to be sorry!* [C13 (in the sense: male servant; C14: young male): of uncertain origin; perhaps from Anglo-French *abuié* fettered (unattested), from Latin *boia* fetter] ■**boy band** *n* an all-male vocal pop group created to appeal to a young audience ■**boyf** (bɔɪf) *n slang* a boyfriend ■**boyfriend** ('bɔɪ,frɛnd) *n* a male friend with whom a person is romantically or sexually involved; sweetheart or lover ■**boyhood** ('bɔɪhʊd) *n* the state or time of being a boy: *his boyhood was happy* ■**boyish** ('bɔɪɪʃ) *adj* of or like a boy in looks, behaviour, or character, esp when regarded as attractive or endearing: *a boyish smile* ▷ **'boyishly** *adv* ▷ **'boyishness** *n* ■**boy-meets-girl** *adj* conventionally or trivially romantic: *a boy-meets-girl story* ■**boy racer** *n informal* **a** a young man who drives his car aggressively and at inappropriately high speeds **b** (*as modifier*): *the boy-racer market* ■**Boys' Brigade** *n* (in Britain) an organization for boys, founded in 1883, with the aim of promoting discipline and self-respect ■**boy scout** *n* **1** See **Scout 2** *US and Canadian informal* an apparently virtuous and innocent person ■**boyshorts** ('bɔɪ,ʃɔːts) *pl n* women's underpants which resemble close-fitting shorts, sitting below the waist and stretching to the tops of the legs ■**boysy** ('bɔɪzɪ) *adj informal* suited to or typical of boys or young men: *done in a matey, boysy way*

boyar ('bəʊjɑː, 'bɔɪə) *n* a member of an old order of Russian nobility, ranking immediately below the princes: abolished by Peter the Great [C16: from Old Russian *boyarin*, from Old Slavonic *boljarinŭ*, probably from Old Turkic *boila* a title] ■**boyarism** ('bɔɪərɪzəm, 'bəʊjɑː,rɪzəm) *n Russian history* the rule of the boyars

boyau ('bɔɪjəʊ) *n fortifications* a minor connecting trench often built in a zigzag pattern

Boyce (bɔɪs) *n* **William.** ?1710–79, English composer, noted esp for his church music and symphonies

boychik ('bɔɪtʃɪk) *n dialect* (esp in Jewish usage) a term of endearment for a boy or young man

boycott ('bɔɪkɒt) *vb* **1** (*tr*) to refuse to have dealings with (a person, organization, etc) or refuse to buy (a product) as a protest or means of coercion: *to boycott foreign produce* ▷ *n* **2** an instance or the use of boycotting [C19: after Captain C. C. Boycott (1832–97), Irish land agent for the Earl of Erne, County Mayo, Ireland, who was a victim of such practices for refusing to reduce rents] ■**boycotter** ('bɔɪkɒtə) *n* a person who boycotts

Boycott ('bɔɪkɒt) *n* **Geoff(rey).** born 1940, English cricketer: played for Yorkshire (1962–86); played in 108 test matches (1964–82); first England batsman to score 8,000 test runs

Boyd (bɔɪd) *n* **1 Arthur.** 1920–99, Australian painter and sculptor, noted for his large ceramic sculptures and his series of engravings **2 Martin** (A'Beckett). 1893–1972, Australian novelist, author of *Lucinda Brayford* (1946) and of the Langton tetralogy *The Cardboard Crown* (1952), *A Difficult Young Man* (1955), *Outbreak of Love* (1957), and *When Blackbirds Sing* (1962) **3 Sir Michael.** born 1955, British theatre director; artistic director of the Royal Shakespeare Company from 2003

Boyd Orr (ɔː) *n* **John, 1st Baron Boyd Orr of Brechin Mearns.** 1880–1971, Scottish biologist; director general of the United Nations Food and Agriculture Organization: Nobel peace prize 1949

Boyer (*French* bwaje) *n* **Charles** (ʃarl), known as *the Great Lover*. 1899–1978, French film actor

boyg (bɔɪɡ) *n Norse myth* a troll-like creature; an ogre

boykie ('bɔɪki:) *n South African informal* a chap or fellow

boyla ('bɔɪlə) *n* an Aboriginal Australian magician or medicine-man

Boyle (bɔɪl) *n* **1 Robert.** 1627–91, Irish scientist who helped to dissociate chemistry from alchemy. He established that air has weight and studied the behaviour of gases; author of *The Sceptical Chymist* (1661) **2 Danny.** born 1956, English film director whose work includes *Trainspotting* (1996) and *Slumdog Millionaire* (2008); artistic director of the opening ceremony of the London 2012 Olympics ■**Boyle's law** *n* the principle that the pressure of a gas varies inversely with its volume at constant temperature [C18: named after Robert Boyle]

Boyne (bɔɪn) *n* a river in the E Republic of Ireland, rising in the Bog of Allen and flowing northeast to the Irish Sea: William III of England defeated the deposed James II in a battle (**Battle of the Boyne**) on its banks in 1690, completing the overthrow of the Stuart cause in Ireland. Length: about 112 km (70 miles)

boyo ('bɔɪəʊ) *n Brit informal* a boy or young man: often used in direct address [from Irish and Welsh]

Boyoma Falls (bɔɪ'əʊmə) *pl n* a series of seven cataracts in the NE Democratic Republic of Congo, on the upper River Congo: forms an unnavigable stretch of 90 km (56 miles), which falls 60 m (200 ft). Former name: **Stanley Falls**

boysenberry ('bɔɪzənbərɪ) *n, pl* **-ries 1** a type of bramble: a hybrid of the loganberry and various blackberries and raspberries **2** the large red edible fruit of this plant [C20: named after Rudolph *Boysen*, American botanist who developed it]

Boz (bɒz) *n* pen name of (Charles) Dickens

Bozcaada (,bɒzdʒaa'da) *n* the Turkish name for **Tenedos**

Bozen ('boːtsən) *n* the German name for **Bolzano**

bozo ('bəʊzəʊ) *n, pl* **-zos** *US slang* a man, esp a stupid one [C20: of uncertain origin; perhaps based on BEAU]

bozzetto *Italian* (bot'zetʊ) *n, pl* **-ti** (-ti) *art* a small model for a planned sculpture or a small sketch for a planned painting

bp *abbreviation for* **1** (of alcoholic density) below proof **2** boiling point **3** bishop

BP *abbreviation for* **1** blood pressure **2** British Pharmacopoeia

BPC *abbreviation for* British Pharmaceutical Codex

B.P.E. *abbreviation for* (in the US and Canada) Bachelor of Physical Education

BPharm *abbreviation for* Bachelor of Pharmacy

BPhil *abbreviation for* Bachelor of Philosophy

bpi *abbreviation for* bits per inch (used of a computer tape or disk surface)

BPR *abbreviation for* business process re-engineering

bps *computing abbreviation for* bits per second (of transmitted information)

b.pt. *abbreviation for* boiling point

Bq *symbol for* becquerel(s)

br *abbreviation for* **1** brother **2** Also: B/R bills receivable ▷ *the internet domain name for* **3** Brazil

Br *abbreviation for* **1** (in a religious order) Brother ▷ *the chemical symbol for* **2** bromine

BR *abbreviation for* **1** (formerly) British Rail **2** Brazil (international car registration)

Br. *abbreviation for* **1** Britain **2** British

B/R *or* **br** *abbreviation for* bills receivable

bra¹ (brɑː) *n* short for **brassiere** ■**braless** ('brɑːlɪs) *adj* not wearing a bra

F

FULL¹, FULFIL] ■ **fillable** ('fɪləbᵊl) *adj* able to be filled ■ **fill away** *vb* (*intr, adverb*) *nautical* to cause a vessel's sails to fill, either by steering it off the wind or by bracing the yards ■ **-filled** *adj* (*in combination*) filled with the specified object: *flower-filled; smoke-filled* ■ **filled gold** *n* another name (esp US) for rolled gold ■ **filler** ('fɪlə) *n* 1 a person or thing that fills 2 an object or substance used to add weight or size to something or to fill in a gap 3 a paste, used for filling in cracks, holes, etc, in a surface before painting 4 *architect* a small joist inserted between and supported by two beams 5 a the inner portion of a cigar b the cut tobacco for making cigarettes 6 *journalism* articles, photographs, etc, to fill space between more important articles in the layout of a newspaper or magazine 7 *informal* something, such as a musical selection, to fill time in a broadcast or stage presentation 8 a small radio or television transmitter used to fill a gap in coverage ■ **filler cap** *n* a device sealing the filling pipe to the petrol tank in a motor vehicle ■ **filler metal** *n* metal supplied in the form of a welding rod, sometimes flux coated, melted by an arc or a flame into a joint between components to be joined ■ **fill in** *vb* (*adverb*) 1 (*tr*) to complete (a form, drawing, etc) 2 (*intr*) to act as a substitute: *a girl is filling in while the typist is away* 3 (*tr*) to put material into (a hole or cavity), esp so as to make it level with a surface 4 (*tr*) *informal* to inform with facts or news 5 (*tr*) *Brit slang* to attack and injure severely ▷ *n* **fill-in** 6 a substitute 7 US *informal* a briefing to complete one's understanding ■ **fill-in** *n* 1 something or someone that temporarily substitutes for something or someone else ▷ *adj* 2 *photog* that supplements the main lighting and e.g. reduces shadows 3 *informal* acting as a temporary substitute for someone or something ■ **filling** ('fɪlɪŋ) *n* 1 the substance or thing used to fill a space or container: *pie filling* 2 *dentistry* a any of various substances (metal, plastic, etc) for inserting into the prepared cavity of a tooth b the cavity of a tooth so filled 3 *textiles* another term for weft ▷ *adj* 4 (of food or a meal) substantial and satisfying ■ **filling station** *n* a place where petrol and other supplies for motorists are sold ■ **fill light** *n* *photog* a light that supplements the key light without changing its character, used esp to lighten shadows ■ **fill out** *vb* (*adverb*) 1 to make or become fuller, thicker, or rounder: *her figure has filled out since her marriage* 2 to make more substantial: *the writers were asked to fill their stories out* 3 (*tr*) to complete (a form, application, etc) ■ **fill up** *vb* (*adverb*) 1 (*tr*) to complete (a form, application, etc) 2 to make or become completely full ▷ *n* **fill-up** 3 the act of filling something completely, esp the petrol tank of a car

fillagree ('fɪləˌgriː) *n, adj, vb* a less common variant of filigree

fille (fiː) *n* a girl or young woman ■ **fille de joie** French (fij də ʒwa) *n, pl* **filles de joie** (fij də ʒwa) a prostitute [girl of pleasure]

fillet ('fɪlɪt) *n* 1 a Also called: **fillet steak** a strip of boneless meat, esp the undercut of a sirloin of beef b the boned side of a fish c the white meat of breast and wing of a chicken 2 a narrow strip of any material 3 a thin strip of ribbon, lace, etc, worn in the hair or around the neck 4 a narrow flat moulding, esp one between other mouldings 5 a narrow band between two adjacent flutings on the shaft of a column 6 Also called: **fillet weld** a narrow strip of welded metal of approximately triangular cross-section used to join steel members at right angles 7 *heraldry* a horizontal division of a shield, one quarter of the depth of the chief 8 Also called: **listel, list** the top member of a cornice 9 *anatomy* a band of sensory nerve fibres in the brain connected to the thalamus. Technical name: **lemniscus** 10 a a narrow decorative line, impressed on the cover of a book b a wheel tool used to impress such lines 11 another name for fairing ▷ *vb* **-lets, -leting, -leted** (*tr*) 12 to cut or prepare (meat or fish) as a fillet 13 to cut fillets from (meat or fish) 14 *anatomy* to surgically remove a bone from (part of the body) so that only soft tissue remains 15 to bind or decorate with or as if with a fillet ▶ Also (for senses 1–3): **filet** [C14: from Old French *filet*, from *fil* thread, from Latin *filum*]

fillip ('fɪlɪp) *n* 1 something that adds stimulation or enjoyment 2 the action of holding a finger towards the palm with the thumb and suddenly releasing it outwards to produce a snapping sound 3 a quick blow or tap made by a finger snapped in this way ▷ *vb* 4 (*tr*) to stimulate or excite 5 (*tr*) to strike or project sharply with a fillip 6 (*intr*) to make a fillip [C15 *philippe*, of imitative origin]

fillipeen (ˌfɪləˈpiːn) *n* another word for philopena

fillister, filister or **fillester** ('fɪlɪstə) *n* 1 Also called: **fillister plane** an adjustable plane for cutting rabbets, grooves, etc 2 Also called: **sash fillister** a rabbet or groove, esp one in a window sash bar for a pane of glass [C19: of unknown origin]

Fillmore ('fɪlmɔː) *n* Millard. 1800–74, 13th president of the US (1850–53); a leader of the Whig Party

filly ('fɪlɪ) *n, pl* **-lies** 1 a female horse or pony under the age of four 2 *informal, rare* a spirited girl or young woman [C15: from Old Norse *fylja*; related to Old High German *fulihha*; see FOAL]

film (fɪlm) *n* 1 a a sequence of images of moving objects photographed by a camera and providing the optical illusion of continuous movement when projected onto a screen b a form of entertainment, information, etc, composed of such a sequence of images and shown in a cinema, etc c (*as modifier*): *film techniques* 2 a thin flexible strip of cellulose coated with a photographic emulsion, used to make negatives and transparencies 3 a thin coating or layer 4 a thin sheet of any material, as of plastic for packaging 5 a fine haze, mist, or blur 6 a gauzy web of filaments or fine threads 7 *pathol* an abnormally opaque tissue, such as the cornea in some

eye diseases ▷ *vb* 8 a to photograph with a cine camera b to make a film of (a screenplay, event, etc) 9 (often foll by *over*) to cover or become covered or coated with a film [Old English *filmen* membrane; related to Old Frisian *filmene*, Greek *pelma* sole of the foot; see FELL⁴] ■ **filmable** ('fɪlməbᵊl) *adj* films able or well-suited to be filmed ■ **filmcard** ('fɪlmˌkɑːd) *n* a cinema loyalty card ■ **film colour** *n* *physiol* a misty appearance produced when no lines or edges are present in the visual field ■ **filmdom** ('fɪlmdəm) *n* *jocular* the cinema industry ■ **filmer** ('fɪlmə) *n* a film-maker ■ **filmgoer** ('fɪlmˌgəʊə) *n* esp Brit a person who goes to see a film or films ■ **filmgoing** ('fɪlmˌgəʊɪŋ) *n* esp Brit the activity of going to see films ■ **filmi** ('fɪlmɪ) *adj* Hinglish 1 of or relating to the Indian film industry or Indian films 2 containing the high drama typical of Indian films [Hindi] ■ **filmic** ('fɪlmɪk) *adj* 1 of or relating to films or the cinema 2 having characteristics that are suggestive of films or the cinema ▷ '**filmically** *adv* ■ **filming** ('fɪlmɪŋ) *n* esp Brit the activity of shooting the scenes of a film ■ **filmish** ('fɪlmɪʃ) *adj* resembling a film ■ **filmland** ('fɪlmˌlænd) *n* Brit jocular the cinema industry ■ **filmless** ('fɪlmləs) *adj* without film ■ **film library** *n* a collection of films as archives or for loan or hire ■ **film-maker** *n* a person who directs or produces films for the cinema or television ■ **film-making** *n* esp Brit the activity or business of producing and directing films ■ **film noir** (nwɑː) *n* a gangster thriller, made esp in the 1940s in Hollywood characterized by contrasty lighting and often somewhat impenetrable plots [C20: French, literally: black film] ■ **filmography** (fɪlˈmɒgrəfɪ) *n* 1 a list of the films made by a particular director, actor, etc 2 any writing that deals with films or the cinema ■ **film pack** *n* a box containing several sheets of film for use in a plate camera ■ **filmset** ('fɪlmˌsɛt) *vb* **-sets, -setting, -set** (*tr*) to set (type matter) by filmsetting ▷ '**film,setter** *n* ■ **film set** *n* the scenery and props as arranged for shooting a film ■ **filmsetting** ('fɪlmˌsɛtɪŋ) *n* printing typesetting by exposing type characters onto photographic film from which printing plates are made ■ **film speed** *n* 1 the sensitivity to light of a photographic film, specified in terms of the film's ISO rating 2 the rate at which the film passes through a motion picture camera or projector ■ **film star** *n* a popular film actor or actress ■ **filmstrip** ('fɪlmˌstrɪp) *n* a strip of film composed of different images projected separately as slides ■ **filmy** ('fɪlmɪ) *adj* **filmier, filmiest** 1 composed of or resembling film; transparent or gauzy 2 covered with or as if with a film; hazy; blurred ▷ '**filmily** *adv* ▷ '**filminess** *n* ■ **filmy fern** *n* any fern of the family *Hymenophyllaceae*, growing in humid regions and having thin translucent leaves

filo ('fiːləʊ) *n* a type of Greek flaky pastry in very thin sheets [C20: Modern Greek *phullon* leaf]

Filofax ('faɪləʊˌfæks) *n trademark* a type of loose-leaf ring binder with sets of different-coloured paper, used as a portable personal filing system, including appointments, addresses, etc

filoplume ('fɪləˌpluːm, -faɪ-) *n ornithol* any of the hairlike feathers that lack vanes and occur between the contour feathers [C19: from New Latin *filoplūma*, from Latin *filum* thread + *plūma* feather]

filopodia (ˌfɪləˈpəʊdɪə) *pl n, sing* **-dium** (-dɪəm) *zoology* thin projections extending from the edge of migrating cells

filoselle (ˌfɪləʊˈsɛl) *n* soft silk thread, used esp for embroidery [C17: from French: silk, silkworm, from Italian *filosello*, perhaps from Latin *folliculus* little bag]

filovirus ('faɪləʊˌvaɪrəs) *n* any member of a family of viruses that includes the agents responsible for Ebola virus disease and Marburg disease [C20: from Latin *filum* thread + VIRUS]

fils¹ French (fis) an addition to a French surname to specify the son rather than the father of the same name: *a book by Dumas fils*. Compare père [French: son]

fils² (fɪls) or **fil** (fɪl) *n, pl* **fils** a a fractional monetary unit of Bahrain, Iraq, Jordan, and Kuwait, worth one thousandth of a dinar b a fractional monetary unit of the United Arab Emirates, worth one hundredth of a dirham c a fractional monetary unit of Yemen, worth one hundredth of a riyal [from Arabic]

filter ('fɪltə) *n* 1 a porous substance, such as paper or sand, that allows fluid to pass but retains suspended solid particles: used to clean fluids or collect solid particles 2 any device containing such a porous substance for separating suspensions from fluids 3 any of various porous substances built into the mouth end of a cigarette or cigar for absorbing impurities such as tar 4 any electronic, optical, or acoustic device that blocks signals or radiations of certain frequencies while allowing others to pass. See also **band-pass filter** 5 any transparent disc of gelatine or glass used to eliminate or reduce the intensity of given frequencies from the light leaving a lamp, entering a camera, etc 6 Brit a traffic signal at a road junction consisting of a green arrow which when illuminated permits vehicles to turn either left or right when the main signals are red ▷ *vb* 7 (often foll by *out*) to remove or separate (suspended particles, wavelengths of radiation, etc) from (a liquid, gas, radiation, etc) by the action of a filter 8 (*tr*) to obtain by filtering 9 (*intr*; foll by *through*) to pass (through a filter or something like a filter): *dust filtered through the screen* 10 (*intr*) to flow slowly; trickle [C16 *filtre* from Medieval Latin *filtrum* piece of felt used as a filter, of Germanic origin; see FELT²] ■ **filterable** ('fɪltərəbᵊl) or **filtrable** ('fɪltrəbᵊl) *adj* 1 capable of being filtered 2 (of most viruses and certain bacteria) capable of passing through the pores of a fine filter ▷ '**filtera'bility** or '**filterableness** *n* ■ **filter bed** *n* 1 a layer of sand or gravel in a tank or reservoir through which a

Illuminati (ɪ,luːmɪˈnɑːtiː) *pl n, sing* **-to** (-təʊ) **1** any of several groups of illuminati, esp in 18th-century France **2** a group of religious enthusiasts of 16th-century Spain who were persecuted by the Inquisition **3** a masonic sect founded in Bavaria in 1778 claiming that the illuminating grace of Christ resided in it alone **4** a rare name for the Rosicrucians

illumine (ɪˈluːmɪn) *vb* a literary word for **illuminate** [C14: from Latin *illūmināre* to make light; see ILLUMINATE] > **il'luminable** *adj* ■ **illuminer** (ɪˈluːmɪnə) *n* an illuminator ■ **illuminism** (ɪˈluːmɪˌnɪzəm) *n* **1** belief in and advocation of special enlightenment **2** the tenets and principles of the Illuminati or of any of several religious or political movements initiated by them > **il'luminist** *n*

illusion (ɪˈluːʒən) *n* **1** a false appearance or deceptive impression of reality: *the mirror gives an illusion of depth* **2** a false or misleading perception or belief; delusion: *he has the illusion that he is really clever* **3** *psychol* a perception that is not true to reality, having been altered subjectively in some way in the mind of the perceiver. See also **hallucination 4** a very fine gauze or tulle used for trimmings, veils, etc [C14: from Latin *illūsiō* deceit, from *illūdere*; see ILLUDE] > **il'lusionary** or **il'lusional** *adj* > **il'lusioned** *adj* ■ **illusionism** (ɪˈluːʒəˌnɪzəm) *n* **1** *philosophy* the doctrine that the external world exists only in illusory sense perceptions **2** the use of highly illusory effects in art or decoration, esp the use of perspective in painting to create an impression of three-dimensional reality ■ **illusionist** (ɪˈluːʒənɪst) *n* **1** a person given to illusions; visionary; dreamer **2** *philosophy* a person who believes in illusionism **3** an artist who practises illusionism **4** a conjuror; magician > **il,lusion'istic** *adj*

illusory (ɪˈluːsərɪ) or **illusive** (ɪˈluːsɪv) *adj* producing, produced by, or based on illusion; deceptive or unreal > **il'lusorily** or **il'lusively** *adv* > **il'lusoriness** or **il'lusiveness** *n*

■ **USAGE** *Illusive* is sometimes wrongly used where *elusive* is meant: *they fought hard, but victory remained elusive (not illusive)*

illust. or **illus.** *abbreviation for* **1** illustrated **2** illustration

illustrate (ˈɪləˌstreɪt) *vb* **1** to clarify or explain by use of examples, analogy, etc **2** (*tr*) to be an example or demonstration of **3** (*tr*) to explain or decorate (a book, text, etc) with pictures **4** (*tr*) an archaic word for **enlighten** [C16: from Latin *illustrāre* to make light, explain, from *lustrāre* to purify, brighten; see LUSTRUM] > **'illus,tratable** *adj* > **'illus,trative** or **'illus,tratory** *adj* > **'illus,tratively** *adv* > **'illus,trator** *n* ■ **illustrated** (ˈɪləstreɪtɪd) *adj* (of a book, text, etc) decorated with or making use of pictures ■ **illustration** (ˌɪləˈstreɪʃən) *n* **1** pictorial matter used to explain or decorate a text **2** an example or demonstration: *an illustration of his ability* **3** the act of illustrating or the state of being illustrated > **,illus'trational** *adj*

illustrious (ɪˈlʌstrɪəs) *adj* **1** of great renown; famous and distinguished **2** glorious or great: *illustrious deeds* **3** *obsolete* shining [C16: from Latin *illustris* bright, distinguished, famous, from *illustrāre* to make light; see ILLUSTRATE] > **il'lustriously** *adv* > **il'lustriousness** *n*

illustrissimo (ˌɪlʊˈstrɪsɪməʊ) *adj* most illustrious (esp as belonging to the Italian aristocracy)

illuviation (ɪ,luːvɪˈeɪʃən) *n* the process by which a material (**illuvium**), which includes colloids and mineral salts, is washed down from one layer of soil to a lower layer [C20: from Latin *illuviēs* dirt, mud, from IL- + -*luviēs*, from *lavere* to wash] > **il'luvial** *adj* ■ **illuviate** (ɪˈluːvɪˌeɪt) *vb* to deposit illuvium

illuvium (ɪˈluːvɪəm) *n geology* a material, which includes colloids and mineral salts, that is washed down from one layer of soil to a lower layer

Illyria (ɪˈlɪərɪə) *n* an ancient region of uncertain boundaries on the E shore of the Adriatic Sea, including parts of present-day Croatia, Montenegro, and Albania ■ **Illyrian** (ɪˈlɪərɪən) *n* **1 a** a member of the group of related Indo-European peoples who occupied Illyria from the late third millennium to the early first millennium BC **2** the extinct and almost unrecorded language of these peoples: of uncertain relationship within the Indo-European family, but thought by some to be the ancestor of modern Albanian ▷ *adj* **3** of, characteristic of, or relating to Illyria, its people, or their language

Illyricum (ɪˈlɪərɪkəm) *n* a Roman province founded after 168 BC, based on the coastal area of Illyria

Ilmen (ˈɪlmən) *n* **Lake Ilmen** a lake in NW Russia, in the Novgorod Region; drains through the Volkhov River into Lake Ladoga. Area: between 780 sq km (300 sq miles) and 2200 sq km (850 sq miles), according to the season ■ **ilmenite** (ˈɪlmɪˌnaɪt) *n* a black mineral found in igneous rocks as layered deposits and in veins. It is the chief source of titanium. Composition: iron titanium oxide. Formula: FeTiO₃. Crystal structure: hexagonal (rhombohedral) [C19: from *Ilmen*, mountain range in the southern Urals, Russia, + -ITE¹]

ILO *abbreviation for* International Labour Organisation

Iloilo (ˌiːləʊˈiːləʊ) *n* a port in the W central Philippines, on SE Panay Island. Pop: 408 000 (2005 est)

Ilorin (ɪˈlɒrɪn) *n* a city in W Nigeria, capital of Kwara state: agricultural trade centre. Pop: 714 000 (2005 est)

ILR *abbreviation for* (in Britain) **1** indefinite leave to remain: an immigration status permitting a person to work or study in the UK without limit of time **2** Independent Local Radio

ILS *aeronautics abbreviation for* instrument landing system

ILU *text messaging abbreviation for* I love you

Ilves (ˈiːlves) *n* **Toomas Hendrik**, born 1953, Estonian politician, president of Estonia from 2006

Ilyushin (Russian ɪlˈjuːʃɪn) *n* **Sergei Vladimirovich** (sɛrˈgei vladiˈmiːrovɪtʃ). 1894–1977, Soviet aircraft designer. He designed the dive bomber Il-2 Stormovik and the jet airliner Il-62

im *the internet domain name for* Isle of Man

IM *abbreviation for* **1** *computing* instant message **2** *computing* instant messaging **3** Also: **i.m.** intramuscular **4** *chess* International Master

im- *prefix* a variant of in-¹, in-²

I'm (aɪm) *contraction of* I am

image (ˈɪmɪdʒ) *n* **1** a representation or likeness of a person or thing, esp in sculpture **2** an optically formed reproduction of an object, such as one formed by a lens or mirror **3** a person or thing that resembles another closely; double or copy **4** a mental representation or picture; idea produced by the imagination **5** the personality presented to the public by a person, organization, etc: *a criminal charge is not good for a politician's image*. See also **corporate image 6** the pattern of light that is focused on to the retina of the eye **7** *psychol* the mental experience of something that is not immediately present to the senses, often involving memory. See also **imagery, body image, hypnagogic image 8** a personification of a specified quality; epitome: *the image of good breeding* **9** a mental picture or association of ideas evoked in a literary work, esp in poetry **10** a figure of speech, such as a simile or metaphor **11** *maths* **a** (of a point) the value of a function, f(x), corresponding to the point x **b** the range of a function **12** an obsolete word for **apparition** ▷ *vb* (*tr*) **13** to picture in the mind; imagine **14** to make or reflect an image of **15** *computing* to project or display on a screen or visual display unit **16** to portray or describe **17** to be an example or epitome of; typify [C13: from Old French *imagene*, from Latin *imāgō* copy, representation; related to Latin *imitārī* to IMITATE] > **'imageable** *adj* > **'imageless** *adj* ■ **image-building** *n* a improving the brand image or public image of something or someone by good public relations, advertising, etc **b** (*as modifier*): *an image-building exercise* ■ **image-conscious** *adj* concerned about the way one comes across to other people and the impression one creates ■ **image converter** or **image tube** *n* a device for producing a visual image formed by other electromagnetic radiation such as infrared or ultraviolet radiation or X-rays ■ **image enhancement** *n* a method of improving the definition of a video picture by a computer program, which reduces the lowest grey values to black and the highest to white: used for pictures from microscopes, surveillance cameras, and scanners ■ **image intensifier** or **image tube** *n* any of various devices for amplifying the intensity of an optical image, sometimes used in conjunction with an image converter ■ **image orthicon** *n* a television camera tube in which electrons, emitted from a photoemissive surface in proportion to the intensity of the incident light, are focused onto the target causing secondary emission of electrons ■ **image printer** *n computing* a printer which uses optical technology to produce an image of a complete page from digital input ■ **image processing** *n* the manipulation or modification of a digitized image, esp in order to enhance its quality ■ **imager** (ˈɪmɪdʒə) *n* an electronic device that records images: *a thermal imager* ■ **imagery** (ˈɪmɪdʒrɪ, -dʒərɪ) *n, pl* **-ries** **1** figurative or descriptive language in a literary work **2** images collectively **3** *psychol* **a** the materials or general processes of the imagination **b** the characteristic kind of mental images formed by a particular individual. See also **image** (7), **imagination** (1) **4** *military* the presentation of objects reproduced photographically (by infrared or electronic means) as prints or electronic displays ■ **image tube** *n* another name for **image converter, image intensifier** ■ **imaging** (ˈɪmɪdʒɪŋ) *n* **1 a** the process of forming or obtaining images by electronically tracing something such as sound waves, temperature, or chemicals, rather than by using light rays or ordinary photography **b** (*as modifier*): *sophisticated imaging technology* **2** *computing* **a** the process of creating images from documents or photographs **b** (*as modifier*): *a printing and imaging business* ■ **imagism** (ˈɪmɪˌdʒɪzəm) *n* a poetic movement in England and America between 1912 and 1917, initiated chiefly by Ezra Pound, advocating the use of ordinary speech and the precise presentation of images > **'imagist** *n, adj* > **,imag'istic** *adj* > **,imag'istically** *adv*

imagine (ɪˈmædʒɪn) *vb* **1** (when *tr, may take a clause as object*) to form a mental image of **2** (when *tr, may take a clause as object*) to think, believe, or guess **3** (*tr; takes a clause as object*) to suppose; assume: *I imagine he'll come* **4** (*tr; takes a clause as object*) to believe or assume without foundation: *he imagines he knows the whole story* **5** an archaic word for **plot¹** ▷ *sentence substitute* **6** Also: **imagine that!** an exclamation of surprise [C14: from Latin *imāgināri* to fancy, picture mentally, from *imāgō* likeness; see IMAGE] > **i'maginable** *adj* > **i'maginableness** *n* > **i'maginably** *adv* > **i'maginer** *n* ■ **imaginal** (ɪˈmædʒɪn²l) *adj* **1** of, relating to, or resembling an imago **2** of or relating to an image ■ **imaginary** (ɪˈmædʒɪnərɪ, -dʒɪnrɪ) *adj* **1** existing in the imagination; unreal; illusory **2** *maths* involving or containing imaginary numbers. The imaginary part of a complex number, z, is usually written Imz > **im'aginarily** *adv* > **im'aginariness** *n* ■ **imaginary number** *n* any complex number of the form ib, where i = √-1 ■ **imaginary part** *n* the coefficient b in a complex number a + ib, where i = √-1 ■ **imagination** (ɪ,mædʒɪˈneɪʃən) *n* **1** the faculty or action of producing ideas, esp mental images of what is not present or has not been experienced **2** mental creative ability **3** the ability to deal resourcefully with unexpected or unusual problems, circumstances, etc **4** (in romantic literary criticism, esp that of S. T. Coleridge) a creative act of perception that joins passive and active elements in thinking and imposes

I

photodissociate (ˌfəʊtəʊdɪˈsəʊsɪeɪt) *vb* (*tr*) *chem* to split or break up molecules as a result of the absorption of photons

photoduplicate (ˌfəʊtəʊˈdjuːplɪkət) *n, vb* the US and Canadian equivalent of photocopy

photodynamics (ˌfəʊtəʊdaɪˈnæmɪks) *n* (*functioning as singular*) the branch of biology concerned with the effects of light on the actions of plants and animals ∎ **photodynamic** (ˌfəʊtəʊdaɪˈnæmɪk) *adj* **1** of or concerned with photodynamics **2** involving or producing an adverse or toxic reaction to light, esp ultraviolet light **3** *med* denoting a therapy for cancer in which a cytotoxic drug is activated by a laser beam

photoelasticity (ˌfəʊtəʊɪlæˈstɪsɪtɪ) *n physics* the effects of stress, such as double refraction, on the optical properties of transparent materials ∎ **photoelastic** (ˌfəʊtəʊɪˈlæstɪk) *adj physics* displaying photoelasticity; of or relating to photoelasticity

photoelectric (ˌfəʊtəʊɪˈlɛktrɪk) or **photoelectrical** *adj* of or concerned with electric or electronic effects caused by light or other electromagnetic radiation ⟩ **photoe'lectrically** *adv* ⟩ **photoelectricity** (ˌfəʊtəʊɪlɛkˈtrɪsɪtɪ) *n* ∎ **photoelectric cell** *n* another name for photocell ∎ **photoelectric effect** *n* **1** the ejection of electrons from a solid by an incident beam of sufficiently energetic electromagnetic radiation **2** any phenomenon involving electricity and electromagnetic radiation, such as photoemission ∎ **photoelectric magnitude** *n astronomy* the magnitude of a star determined using a photometer plus a filter to select light or other radiation of the desired wavelength

photoelectrode (ˌfəʊtəʊɪˈlɛktrəʊd) *n physics* an electrode that, following the absorption of light, can initialize electrochemical transformations

photoelectron (ˌfəʊtəʊɪˈlɛktrɒn) *n* an electron ejected from an atom, molecule, or solid by an incident photon ∎ **photoelectronic** (ˌfəʊtəʊˌiːlɛkˈtrɒnɪk, ˌfəʊtəʊˌɪlɛkˈtrɒnɪk) *adj physics* **1** relating to electronic effects or devices affected by light **2** of or relating to a photoelectron

photoelectrotype (ˌfəʊtəʊɪˈlɛktrəʊˌtaɪp) *n* an electrotype mode using photography

photoemission (ˌfəʊtəʊɪˈmɪʃən) *n* the emission of electrons due to the impact of electromagnetic radiation, esp as a result of the photoelectric effect ⟩ **photoe'missive** *adj*

photoengraving (ˌfəʊtəʊɪnˈɡreɪvɪŋ) *n* **1** a photomechanical process for producing letterpress printing plates **2** a plate made by this process **3** a print made from such a plate ∎ **photoengrave** (ˌfəʊtəʊɪnˈɡreɪv) *vb* (*tr*) to reproduce (an illustration) by photoengraving ⟩ **photoen'graver** *n*

photoexcitation (ˌfəʊtəʊˌɛksɪˈteɪʃən) *n chem, physics* the creation of an increase in energy in atoms, molecules or ions caused by the absorption of a photon ∎ **photoexcited** (ˌfəʊtəʊɪkˈsaɪtɪd) *adj chem, physics* increased in energy due to photoexcitation

photofission (ˌfəʊtəʊˈfɪʃən) *n physics* the simultaneous release of energy and splitting of a large nucleus into smaller ones, caused by the absorption of a high energy photon

Photofit (ˈfəʊtəʊˌfɪt) *n trademark* **a** a method of combining photographs of facial features, hair, etc, into a composite picture of a face: formerly used by the police to trace suspects from witnesses' descriptions **b** (*as modifier*): *a Photofit picture*

photofluorogram (ˌfəʊtəʊˈflʊərəˌɡræm) *n med, photog* a photographic image produced by photofluorography

photofluorography (ˌfəʊtəʊflʊəˈrɒɡrəfɪ) *n med* the process of taking a photograph (**photofluorogram**) of a fluoroscopic image: used in diagnostic screening

photog. *abbreviation for* **1** photograph **2** photographer **3** photographic **4** photography

photogelatine (ˌfəʊtəʊˈdʒɛlɪtiːn) *adj printing* relating to photographic processes in which gelatine is used in receiving or transferring prints ∎ **photogelatine process** (ˌfəʊtəʊˈdʒɛlətiːn) *n* another name for **collotype** (1)

photogelatin process *n* another name for photogelatine process

photogene (ˈfəʊtəʊˌdʒiːn) *n* another name for afterimage [C19: from Greek *phōtogenēs* light-produced. See PHOTO-, -GENE]

photogenic (ˌfəʊtəʊˈdʒɛnɪk) *adj* **1** (esp of a person) having features, colouring, and a general facial appearance that look attractive in photographs **2** *biology* producing or emitting light: *photogenic bacteria* ⟩ **photo'genically** *adv*

photogeny (fəˈtɒdʒɪnɪ) *n* an obsolete name for photography

photogeology (ˌfəʊtəʊdʒɪˈɒlədʒɪ) *n* the study and identification of geological phenomena using aerial photographs ∎ **photogeologic** (ˌfəʊtəʊˌdʒɪəˈlɒdʒɪk) or **photogeological** (ˌfəʊtəʊˌdʒɪəˈlɒdʒɪkəl) *adj* of or relating to photogeology ∎ **photogeologist** (ˌfəʊtəʊdʒɪˈɒlədʒɪst) *n* a person who studies or has a profession in photogeology

photoglyph (ˈfəʊtəʊˌɡlɪf) *n* an engraved plate, produced by the action of light, and from which prints or impressions are taken ∎ **photoglyphic** (ˌfəʊtəʊˈɡlɪfɪk) *adj* of or relating to photoglyphy ∎ **photoglyphy** (ˈfəʊtəʊˌɡlɪfɪ) *n* the art or process of engraving using the action of light

photogram (ˈfəʊtəʊˌɡræm) *n* **1** a picture, usually abstract, produced on a photographic material without the use of a camera, as by placing an object on the material and exposing to light **2** *obsolete* a photograph, often of the more artistic kind rather than a mechanical record ∎ **photogrammetry** (ˌfəʊtəʊˈɡræmɪtrɪ) *n* the process of making measurements from photographs, used esp in the construction of maps from aerial photographs and also in military intelligence, medical and industrial research, etc ⟩ **photogrammetric** (ˌfəʊtəʊɡrəˈmɛtrɪk) *adj* ⟩ **photo'grammetrist** *n*

photograph (ˈfəʊtəˌɡrɑːf, -ˌɡræf) *n* **1** an image of an object, person, scene, etc, in the form of a print or slide recorded by a camera on photosensitive material. Often shortened to: **photo** ▷ *vb* **2** to take a photograph of (an object, person, scene, etc) ∎ **photographer** (fəˈtɒɡrəfə) *n* a person who takes photographs, either as a hobby or a profession ∎ **photographic** (ˌfəʊtəˈɡræfɪk) or **photographical** (ˌfəʊtəˈɡræfɪkəl) *adj* **1** of or relating to photography: *a photographic society; photographic materials* **2** like a photograph, in accuracy or detail **3** (of a person's memory) able to retain facts, appearances, etc, in precise detail, often after only a very short view of or exposure to them ⟩ **photo'graphically** *adv* ∎ **photographist** (fəˈtɒɡrəfɪst) *n* another word for **photographer** ∎ **photography** (fəˈtɒɡrəfɪ) *n* **1** the process of recording images on sensitized material by the action of light, X-rays, etc, and the chemical processing of this material to produce a print, slide, or cine film **2** the art, practice, or occupation of taking and printing photographs, making cine films, etc

photogravure (ˌfəʊtəʊɡrəˈvjʊə) *n* **1** any of various methods in which an intaglio plate for printing is produced by the use of photography **2** matter printed from such a plate ▸ Former name: **heliogravure** [C19: from PHOTO- + French *gravure* engraving]

photoinduced (ˌfəʊtəʊɪnˈdjuːst) *adj* induced by exposure to light or other electromagnetic radiation ∎ **photoinduction** (ˌfəʊtəʊɪnˈdʌkʃən) *n* the act or process of being photoinduced ∎ **photoinductive** (ˌfəʊtəʊɪnˈdʌktɪv) *adj* of, relating to, or being able to undergo photoinduction

photoionize or **photoionise** (ˌfəʊtəʊˈaɪəˌnaɪz) *vb* (*tr*) *physics* to cause to undergo or to undergo photoionization ∎ **photoionization** or **photoionisation** (ˌfəʊtəʊˌaɪənaɪˈzeɪʃən) *n physics* the ejection of electrons from an atom, ion or molecule caused by the absorption of photons

photokinesis (ˌfəʊtəʊkɪˈniːsɪs, -kaɪ-) *n biology* the movement of an organism in response to the stimulus of light ⟩ **photokinetic** (ˌfəʊtəʊkɪˈnɛtɪk, -kaɪ-) *adj* ⟩ **photoki'netically** *adv*

photolitho (ˌfəʊtəʊˈlaɪθəʊ) *printing n* **1** photolithography **2** (*pl* **-thos**) photolithograph ▷ *adj* **3** photolithographic

photolithograph (ˌfəʊtəʊˈlɪθəˌɡrɑːf, -ˌɡræf) *n* **1** a picture printed by photolithography ▷ *vb* **2** (*tr*) to reproduce (pictures, text, etc) by photolithography ∎ **photolithography** (ˌfəʊtəʊlɪˈθɒɡrəfɪ) *n* **1** a lithographic printing process using photographically made plates. Often shortened to: **photolitho** (ˌfəʊtəʊˈlaɪθəʊ) **2** *electronics* a process used in the manufacture of semiconductor devices, thin-film circuits, optical devices, and printed circuits in which a particular pattern is transferred from a photograph onto a substrate, producing a pattern that acts as a mask during an etching or diffusion process. See also planar process ⟩ **photoli'thographer** *n* ⟩ **photolithographic** (ˌfəʊtəʊˌlɪθəˈɡræfɪk) *adj* ⟩ **photo₁litho'graphically** *adv*

photoluminesce (ˌfəʊtəʊˌluːmɪˈnɛs) *vb* (*tr*) to produce photoluminescence ∎ **photoluminescence** (ˌfəʊtəʊˌluːmɪˈnɛsəns) *n* luminescence resulting from the absorption of light or infrared or ultraviolet radiation ⟩ **photo₁lumi'nescent** *adj*

photolysis (fəʊˈtɒlɪsɪs) *n* chemical decomposition caused by light or other electromagnetic radiation. Compare radiolysis ⟩ **photolytic** (ˌfəʊtəʊˈlɪtɪk) *adj* ∎ **photolyse** (ˈfəʊtəˌlaɪz) *vb* (*tr*) *chem* to cause to undergo or to undergo photolysis ∎ **photolytically** (ˌfəʊtəʊˈlɪtɪklɪ) *adv chem* in a photolytic manner, by photolytic means ∎ **photolyzable** or **photolysable** (ˈfəʊtəˌlaɪzəbəl) *adj* able to undergo photolysis

photomacrograph (ˌfəʊtəʊˈmækrəʊˌɡrɑːf) *n* **1** *photog* a photograph which shows an object at its actual size or slightly magnified **2** *microscopy* an image made using a low powered microscope

photomask (ˈfəʊtəʊˌmɑːsk) *n* an opaque image on a transparent plate that is used to filter light so the image can be transferred, used in photolithography application

photomechanical (ˌfəʊtəʊmɪˈkænɪkəl) *adj* **1** of or relating to any of various methods by which printing plates are made using photography ▷ *n* **2** a final paste-up of artwork or typeset matter or both for photographing and processing into a printing plate ▸ Often shortened to: **mechanical** ⟩ **photome'chanically** *adv* ∎ **photomechanical transfer** *n* a method of producing photographic prints or offset printing plates from paper negatives by a chemical transfer process rather than by exposure to light

photometer (fəʊˈtɒmɪtə) *n* an instrument used in photometry, usually one that compares the illumination produced by a particular light source with that produced by a standard source. See also spectrophotometer ∎ **photometry** (fəʊˈtɒmɪtrɪ) *n* **1** the measurement of the intensity of light **2** the branch of physics concerned with such measurements ⟩ **photometric** (ˌfəʊtəˈmɛtrɪk) *adj* ⟩ **photo'metrically** *adv* ⟩ **pho'tometrist** *n*

photomicrograph (ˌfəʊtəʊˈmaɪkrəˌɡrɑːf, -ˌɡræf) *n* **1** a photograph of a microscope image. Sometimes called: **microphotograph** **2** a less common name for microphotograph (1) ⟩ **photomicrographer** (ˌfəʊtəʊmaɪˈkrɒɡrəfə) *n* ⟩ **photomicrographic** (ˌfəʊtəʊˌmaɪkrəˈɡræfɪk) *adj* ⟩ **photo₁micro'graphically** *adv* ⟩ **photomi'crography** *n*

photomultiplier (ˌfəʊtəʊˈmʌltɪˌplaɪə) *n* a device sensitive to electromagnetic radiation, consisting of a photocathode, from which electrons are released by incident photons, and an electron multiplier, which amplifies and produces a detectable pulse of current

P

vicious wind **4** characterized by malice: *vicious lies* **5** (esp of dogs, horses, etc) ferocious or hostile; dangerous **6** characterized by or leading to vice **7** invalidated by defects; unsound: *a vicious inference* **8** *obsolete* noxious or morbid: *a vicious exhalation* [C14: from Old French *vicieus*, from Latin *vitiōsus* full of faults, from *vitium* a defect] > **'viciously** *adv* > **'viciousness** *n* ■ **vicious circle** *n* **1** Also: **vicious cycle** a situation in which an attempt to resolve one problem creates new problems that lead back to the original situation **2** *logic* **a** a course of reasoning in which a conclusion is inferred from premises the truth of which cannot be established independently of that conclusion **b** an explanation given in terms that cannot be understood independently of that which was to be explained **c** a situation in which some statement is shown to entail its negation and vice versa, as *this statement is false* is true only if false and false only if true **3** *med* a condition in which one disease or disorder causes another, which in turn aggravates the first condition

vicissitude (vɪˈsɪsɪˌtjuːd) *n* **1** variation or mutability in nature or life, esp successive alternation from one condition or thing to another **2** a variation in circumstance, fortune, character, etc [C16: from Latin *vicissitūdō*, from *vicis* change, alternation] > **vi,cissiˈtudinary** *or* **vi,cissiˈtudinous** *adj*

Vicksburg (ˈvɪksˌbɜːɡ) *n* a city in W Mississippi, on the Mississippi River: site of one of the most decisive campaigns (1863) of the American Civil War, in which the Confederates were besieged for nearly seven weeks before capitulating. Pop: 26 005 (2003 est)

Vicky (ˈvɪkɪ) *n* professional name of *Victor Weisz*. 1913–66, British left-wing political cartoonist, born in Germany

Vico (ˈviːkəʊ; *Italian* ˈviːko) *n* **Giovanni Battista** (dʒoˈvanni batˈtista). 1668–1744, Italian philosopher. In *Scienza Nuova* (1721) he postulated that civilizations rise and fall in evolutionary cycles, making use of myths, poetry, and linguistics as historical evidence

vicomte (*French* vikɔ̃t) *or feminine* **vicomtesse** (*French* vikɔ̃tɛs) *n* a French noble holding a rank corresponding to that of a British viscount or viscountess

Victa (ˈvɪktə) *n trademark Austral* a type of rotary lawnmower first manufactured in 1952 [C20: named after Mervyn *Victor* Richardson, who invented it]

victim (ˈvɪktɪm) *n* **1** a person or thing that suffers harm, death, etc, from another or from some adverse act, circumstance, etc: *victims of tyranny* **2** a person who is tricked or swindled; dupe **3** a living person or animal sacrificed in a religious rite [C15: from Latin *victima*] ■ **victimhood** (ˈvɪktɪmhʊd) *n* the state of being a victim ■ **victimize** *or* **victimise** (ˈvɪktɪˌmaɪz) *vb* (*tr*) **1** to punish or discriminate against selectively or unfairly **2** to make a victim of **3** to kill as or in a manner resembling a sacrificial victim > **,victimiˈzation** *or* **,victimiˈsation** *n* > **'victim,izer** *or* **'victim,iser** *n* ■ **victimless crime** (ˈvɪktɪmlɪs) *n* a type of crime, such as insurance fraud, regarded by some people as being excusable because the victim is the state or an organization, rather than an individual ■ **victimology** (ˌvɪktɪˈmɒlədʒɪ) *n* the study of the psychological effects experienced by the victims of crime > **,victiˈmologist** *n*

● **USAGE** Using the word *victim* or *victims* in relation to chronic illness or disability is often considered demeaning and disempowering. Alternative phrases such as *who experiences, who has been diagnosed with,* or simply *with* and then the name of the disability or illness, can be used instead

victor (ˈvɪktə) *n* **1 a** a person, nation, etc, that has defeated an adversary in war, etc **b** (*as modifier*): *the victor army* **2** the winner of any contest, conflict, or struggle [C14: from Latin, from *vincere* to conquer]

Victor (ˈvɪktə) *n communications* a code word for the letter v

Victor Emmanuel II *n* 1820–78, king of Sardinia-Piedmont (1849–78) and first king of Italy from 1861

Victor Emmanuel III *n* 1869–1947, last king of Italy (1900–46): dominated after 1922 by Mussolini, whom he appointed as premier; abdicated

victoria (vɪkˈtɔːrɪə) *n* **1** a light four-wheeled horse-drawn carriage with a folding hood, two passenger seats, and a seat in front for the driver **2** Also called: **victoria plum** *Brit* a large sweet variety of plum, red and yellow in colour **3** any South American giant water lily of the genus *Victoria*, having very large floating leaves and large white, red, or pink fragrant flowers: family Nymphaeaceae [C19: all named after Queen VICTORIA]

Victoria¹ (vɪkˈtɔːrɪə) *n* **1 a** a state of SE Australia: part of New South Wales colony until 1851; semiarid in the northwest, with the Great Dividing Range in the centre and east and the Murray River along the N border. Capital: Melbourne. Pop: 5 713 000 (2013 est). Area: 227 620 sq km (87 884 sq miles) **2 Lake Victoria** *or* **Victoria Nyanza** a lake in East Africa, in Tanzania, Uganda, and Kenya, at an altitude of 1134 m (3720 ft): the largest lake in Africa and second largest in the world; drained by the Victoria Nile. Area: 69 485 sq km (26 828 sq miles) **3** a port in SW Canada, capital of British Columbia, on Vancouver Island: founded in 1843 by the Hudson's Bay Company; made capital of British Columbia in 1868; university (1963). Pop: 80 032 (2011) **4** the capital of the Seychelles, a port on NE Mahé. Pop: 25 500 (2004 est) **5** an urban area in S China, part of Hong Kong, on N Hong Kong Island: financial and administrative district; university (1911); the name tends not to be used officially since reunification of Hong Kong with China in 1997 **6 Mount Victoria** a mountain in SE Papua New Guinea: the highest peak of the Owen Stanley Range. Height: 4073 m (13 363 ft)

■ **Victoria Desert** *n* See Great Victoria Desert ■ **Victoria Falls** *pl n* a major waterfall on the border between Zimbabwe and Zambia, on the Zambezi River. Height: about 108 m (355 ft). Width: about 1400 m (4500 ft). Local name: **Mosi-oa-Tunya**

Victoria² (vɪkˈtɔːrɪə) *n* **1** 1819–1901, queen of the United Kingdom (1837–1901) and empress of India (1876–1901). She married Prince Albert of Saxe-Coburg-Gotha (1840). Her sense of vocation did much to restore the prestige of the British monarchy **2** (*Spanish* bikˈtorja) **Tomás Luis de**. ?1548–1611, Spanish composer of motets and masses in the polyphonic style ■ **Victoria and Albert Museum** *n* a museum of the fine and applied arts in London, originating from 1856 and given its present name and site in 1899. Abbreviation: **V and A** ■ **Victoria Cross** *n* the highest decoration for gallantry in the face of the enemy awarded to the British and Commonwealth armed forces: instituted in 1856 by Queen Victoria ■ **Victoria Day** *n* the Monday preceding May 24: observed in Canada as a national holiday in commemoration of the birthday of Queen Victoria ■ **Victorian** (vɪkˈtɔːrɪən) *adj* **1** of, relating to, or characteristic of Queen Victoria or the period of her reign **2** exhibiting the characteristics popularly attributed to the Victorians, esp prudery, bigotry, or hypocrisy. Compare **Victorian values** **3** denoting, relating to, or having the style of architecture used in Britain during the reign of Queen Victoria, characterized by massive construction and elaborate ornamentation **4** of or relating to Victoria (the state or any of the cities) > *n* **5** a person who lived during the reign of Queen Victoria **6** an inhabitant of Victoria (the state or any of the cities) > **Vicˈtorian,ism** *n* ■ **Victoriana** (vɪkˌtɔːrɪˈɑːnə) *pl n* objects, ornaments, etc, of the Victorian period ■ **Victorian values** *pl n* qualities considered to characterize the Victorian period, including enterprise and initiative and the importance of the family. Compare Victorian (2)

Victoria³ (vɪkˈtɔːrɪə) *n* the Roman goddess of victory. Greek counterpart: Nike

Victoria Island *n* a large island in the Canadian Arctic, in Nunavut and the Northwest Territories. Area: about 212 000 sq km (82 000 sq miles)

Victoria Land *n* a section of Antarctica, largely in the Ross Dependency on the Ross Sea

Victoria Nile *n* See Nile

victorine (ˈvɪktəriːn) *n* a woman's fur shoulder cape, which fastens at the back

victory (ˈvɪktərɪ) *n*, *pl* **-ries** **1** final and complete superiority in a war **2** a successful military engagement **3** a success attained in a contest or struggle or over an opponent, obstacle, or problem **4** the act of triumphing or state of having triumphed [C14: from Old French *victorie*, from Latin *victōria*, from *vincere* to subdue] ■ **victorious** (vɪkˈtɔːrɪəs) *adj* **1** having defeated an adversary: *the victorious nations* **2** of, relating to, indicative of, or characterized by victory: *a victorious conclusion* > **vicˈtoriously** *adv* > **vicˈtoriousness** *n* ■ **Victory** (ˈvɪktərɪ) *n* **1** another name (in English) for Victoria³ **2** another name (in English) for Nike ■ **victory lap** *n* another name for lap of honour ■ **victoryless** (ˈvɪktərɪlɪs) *adj* without victory ■ **victory roll** *n* a roll of an aircraft made by a pilot to announce or celebrate the shooting down of an enemy plane or other cause for celebration

victress (ˈvɪktrɪs) *n* a female victor

victrola (vɪkˈtrəʊlə) *n* a gramophone

victual (ˈvɪtʲl) *vb* **-uals, -ualling, -ualled** *or US* **-uals, -ualing, -ualed** **1** to supply with or obtain victuals **2** (*intr*) *rare* (esp of animals) to partake of victuals [C14: from Old French *vitaille*, from Late Latin *victuālia* provisions, from Latin *victuālis* concerning food, from *victus* sustenance, from *vīvere* to live] > **'victual-less** *adj* ■ **victualer** (ˈvɪtʲlə) *n* **1** a provider of food or provisions **2** *Brit* an innkeeper **3** a ship with supplies ■ **victuallage** *or* **victualage** (ˈvɪtəlɪdʒ) *n* a rare word for victuals ■ **victualler** (ˈvɪtələ, ˈvɪtlə) *n* **1** a supplier of victuals, as to an army; sutler **2** *Brit* a licensed purveyor of spirits; innkeeper **3** a supply ship, esp one carrying foodstuffs ■ **victuals** (ˈvɪtʲlz) *pl n* (*sometimes singular*) food or provisions

vicuña (vɪˈkuːnjə) *or* **vicuna** (vɪˈkjuːnə) *n* **1** a tawny-coloured cud-chewing Andean artiodactyl mammal, *Vicugna vicugna*, similar to the llama: family Camelidae **2** the fine light cloth made from the wool obtained from this animal [C17: from Spanish *vicuña*, from Quechuan *wikúña*]

vid (vɪd) *n informal* short for video (4)

Vidal (vɪˈdæl) *n* **Gore**. 1925–2012 US novelist and essayist. His novels include *Julian* (1964), *Myra Breckinridge* (1968), *Burr* (1974), *Lincoln* (1984), and *The Season of Conflict* (1996)

vidame (ˈviːdaːm) *n* a French feudal nobleman

vide (ˈvaɪdɪ) *vb* (used to direct a reader to a specified place in a text, another book, etc) refer to; see (often in the phrases **vide ante** (see before), **vide infra** (see below), **vide post** (see after), **vide supra** (see above), **vide ut supra** (see as above), etc). Abbreviation: **v, vid** [C16: from Latin]

videlicet (vɪˈdiːlɪˌsɛt) *adv* namely: used to specify items, examples, etc. Abbreviation: **viz** [C15: from Latin]

videndum (vɪˈdɛndʊm, vaɪ-) *n* the thing which is to be seen [Latin]

video (ˈvɪdɪˌəʊ) *adj* **1** relating to or employed in the transmission or reception of a televised image **2** of, concerned with, or operating at video frequencies > *n*, *pl* **-os** **3** the visual elements of a television broadcast **4** a film recorded on a video cassette **5** short for **video cassette**, **video cassette recorder** **6** *US* an informal name for **television** > *vb* **videos, videoing,**

Case 1:22-cv-00647-WCB

2205 | vidette – view

videoed 7 to record (a television programme, etc) on a video cassette recorder ▸ Compare **audio** [C20: from Latin *vidēre* to see, on the model of AUDIO] ■ **video call** n a call made via a mobile phone with a camera and a screen, allowing the participants to see each other as they talk ■ **videocam** ('vɪdɪəʊ,kæm) n a portable camera that records moving images ■ **video cassette** or **videocassette** ('vɪdɪəʊkə,sɛt) n a cassette containing video tape ■ **video cassette recorder** n a tape recorder for vision and sound signals using magnetic tape in closed plastic cassettes: used for recording and playing back television programmes and films. Often shortened to: **video**. Abbreviation: VCR ■ **videoconference** ('vɪdɪəʊ,kɒnfərəns) n a conference in which participants in distant locations to take part by means of electronic sound and video communication ■ **videodisk** ('vɪdɪəʊ,dɪsk) n another name for optical disc ■ **videofit** ('vɪdɪəʊ,fɪt) n a computer-generated picture of a person sought by the police, created by combining facial characteristics on the basis of witnesses' descriptions [C20: from VIDEO + (PHOTO)FIT] ■ **video frequency** n the frequency of a signal conveying the image and synchronizing pulses in a television broadcasting system. It lies in the range from about 50 hertz to 8 megahertz ■ **video game** n any of various games that can be played by using an electronic control to move points of light or graphical symbols on the screen of a visual display unit ■ **videogram** ('vɪdɪəʊ,græm) n an audiovisual recording, as on a videotape or DVD ■ **videography** (,vɪdɪ'ɒɡrəfɪ) n the art, practice, or occupation of making videos ▸ **,vide'ographer** n ■ **video jockey** n a person who introduces and plays videos, esp of pop songs, on a television programme. Abbreviation: VJ ■ **videoland** ('vɪdɪəʊ,lænd) n the world of television and televised images ■ **video memory** n computing computer memory used for the processing and displaying of images ■ **video nasty** n a film, usually specially made for video, that is explicitly horrific, brutal, and pornographic ■ **videophile** ('vɪdɪəʊ,faɪl) n someone who is interested in videos, in particular with regards to features such as high definition, audio quality, etc ■ **videophone** ('vɪdɪə,fəʊn) or **videotelephone** (,vɪdɪəʊ'tɛlɪ,fəʊn) n a telephonic device in which there is both verbal and visual communication between parties. Also called: **viewphone** ▸ **videophonic** (,vɪdɪə'fɒnɪk) adj ■ **video referee** n rugby an additional referee during a televised game who is able to examine video playback to determine whether or not a try has been legitimately scored ■ **video tape** or **videotape** n 1 magnetic tape used mainly for recording the vision and sound signals of a television programme or film for subsequent transmission ▸ vb **video-tape** 2 to record (a programme, film, etc) on video tape ■ **video tape recorder** n a tape recorder for vision signals and sometimes accompanying sound, using magnetic tape on open spools: used in television broadcasting. Abbreviation: VTR ■ **Videotex** ('vɪdɪəʊ,tɛks) n trademark an information system that displays information from a distant computer on a television screen. See also **Teletext, Viewdata** ■ **videotext** ('vɪdɪəʊ,tɛkst) n a means of providing a written or graphical representation of computerized information on a television screen ■ **videotheque** ('vɪdɪəʊtɛk) n a cinema in which videos are shown

vidette (vɪ'dɛt) n a variant spelling of **vedette**
Vidhan Sabha (vɪ'dɑːn 'sʌbə) n the legislative assembly of any of the states of India [Hindi, from *vidhan* law + *sabha* assembly]
vidicon ('vɪdɪ,kɒn) n a small television camera tube, used in closed-circuit television and outside broadcasts, in which incident light forms an electric charge pattern on a photoconductive surface. Scanning by a low-velocity electron beam discharges the surface, producing a current in an adjacent conducting layer. See also **Plumbicon** [C20: from VID(EO) + ICON(OSCOPE)]
vidimus ('vaɪdɪmʊs, 'viːdɪmʊs) n a legal or official inspection
viduage ('vɪdjʊɪdʒ) n obsolete widows collectively; widowhood
vidual ('vɪdjʊəl) adj formal widowed
viduity (vɪ'djuːɪtɪ) n formal widowhood
viduous ('vɪdjʊəs) adj 1 formal empty 2 a variant form of **vidual**
vie (vaɪ) vb **vies, vying, vied** 1 (intr; foll by with or for) to contend for superiority or victory (with) or strive in competition (for) 2 (tr) archaic to offer, exchange, (with) or display in rivalry [C15: probably from Old French *envier* to challenge, from Latin *invītāre* to INVITE] ▸ **'vier** n ▸ **'vying** adj, n
vielle (viː'ɛl) n a European stringed musical instrument from Medieval times and somewhat similar to a violin
Vienna (vɪ'ɛnə) n the capital and the smallest state of Austria, in the northeast on the River Danube: seat of the Hapsburgs (1278-1918); residence of the Holy Roman Emperor (1558-1806); withstood sieges by Turks in 1529 and 1683; political and cultural centre in the 18th and 19th centuries, having associations with many composers; university (1365). Pop: 1 590 242 (2003 est). Area: 1075 sq km (415 sq miles). German name: **Wien**. Latin name: **Vindobona** ■ **Vienna Union** or **Vienna International** n the Vienna Union an international conference of socialists who came together in Vienna in 1921 in an attempt to reconstruct a united International by offering an alternative to the right-wing remnant of the Second International and to the Comintern: merged into the Labour and Socialist International in 1923. Also known as: **the Two-and-a-half International** ■ **Viennese** (,vɪə'niːz) adj 1 of, relating to, or characteristic of Vienna ▸ n, pl **-nese** 2 a native or inhabitant of Vienna
Vienne (French vjɛn) n 1 a department of W central France, in Poitou-

Charentes region. Capital: Poitiers. Pop: 402 555 (2003 est). Area: 7044 sq km (2747 sq miles) 2 a town in SE France, on the River Rhône: extensive Roman remains. Ancient name: **Vienna** 3 a river in SW central France, flowing west and north to the Loire below Chinon. Length: over 350 km (200 miles)
Vientiane (,vjɛntɪ'ɑːn) or **Viangchan** (wiːɛŋtæn) n the administrative capital of Laos, in the south near the border with Thailand: capital of the kingdom of Vientiane from 1707 until taken by the Thais in 1827. Pop: 776 000 (2005 est)
Vierwaldstättersee (fiːr'valtʃtɛtər,zeː) n the German name for (Lake) Lucerne
vies (fiːs) adj South African slang angry, furious, or disgusted [Afrikaans]
vi et armis Latin (vaɪ ɛt 'ɑːmɪs) n legal history a kind of trespass accompanied by force and violence [literally: by force and arms]
Vietcong (,vjɛt'kɒŋ) or **Viet Cong** n (in the Vietnam War) 1 the Communist-led guerrilla force and revolutionary army of South Vietnam; the armed forces of the National Liberation Front of South Vietnam 2 a member of these armed forces 3 (modifier) of or relating to the Vietcong or a Vietcong [from Vietnamese *Viet Nam Cong San* Vietnamese Communist]
Vietminh (,vjɛt'mɪn) or **Viet Minh** n 1 a Vietnamese organization led by Ho Chi Minh that first fought the Japanese and then the French (1941-54) in their attempt to achieve national independence 2 a member or group of members of this organization, esp in the armed forces 3 (modifier) of or relating to this organization or to its members [from Vietnamese *Viet Nam Doc Lap Dong Minh Hoi* Vietnam League of Independence]
Vietnam (,vjɛt'næm) or **Viet Nam** n a republic in SE Asia: an ancient empire, conquered by France in the 19th century; occupied by Japan (1940-45) when the Communist-led Vietminh began resistance operations that were continued against restored French rule after 1945. In 1954 the country was divided along the 17th parallel, establishing North Vietnam (under the Vietminh) and South Vietnam (under French control), the latter becoming the independent **Republic of Vietnam** in 1955. From 1959 the country was dominated by war between the Communist Vietcong, supported by North Vietnam, and the South Vietnamese government; increasing numbers of US forces were brought to the aid of the South Vietnamese army until a peace agreement (1973) led to the withdrawal of US troops; further fighting led to the eventual defeat of the South Vietnamese government in March 1975 and in 1976 an elected National Assembly proclaimed the reunification of the country. Official language: Vietnamese. Religion: Buddhist majority. Currency: dong. Capital: Hanoi. Pop: 92 477 857 (2013 est). Area: 331 041 sq km (127 816 sq miles). Official name: **Socialist Republic of Vietnam** ■ **Vietnamese** (,vjɛtnə'miːz) adj 1 of, relating to, or characteristic of Vietnam, its people, or their language ▸ n 2 (pl -ese) a native or inhabitant of Vietnam 3 the language of Vietnam, probably related to the Mon-Khmer languages
Vietnamization or **Vietnamisation** (,vjɛtnəmaɪ'zeɪʃən) n (in the Vietnam War) a US government policy of transferring the tasks of fighting and directing the war to the government and forces of South Vietnam
vieux jeu French (vjø ʒø) adj old-fashioned [literally: old game]
view (vjuː) n 1 the act of seeing or observing; an inspection 2 vision or sight, esp range of vision: *the church is out of view* 3 a scene, esp of a fine tract of countryside: *the view from the top was superb* 4 a pictorial representation of a scene, such as a photograph 5 (sometimes plural) opinion; thought: *my own view on the matter differs from yours* 6 chance or expectation: *the policy has little view of success* 7 (foll by to) a desired end or intention: *he has a view to securing further qualifications* 8 a general survey of a topic, subject, etc: *a comprehensive view of Shakespearean literature* 9 visual aspect or appearance: *they look the same in outward view* 10 law a a formal inspection by a jury of the place where an alleged crime was committed b a formal inspection of property in dispute 11 a sight of a hunted animal before or during the chase 12 **in view of** taking into consideration 13 **on view** exhibited to the public gaze 14 **take a dim view of** or **take a poor view of** to regard (something) with disfavour or disapproval 15 **with a view to** a with the intention of b in anticipation or hope of ▸ vb 16 (tr) to look at 17 (tr) to consider in a specified manner: *they view the growth of Communism with horror* 18 (tr) to examine or inspect carefully: *to view the accounts* 19 (tr) to survey mentally; contemplate: *to view the difficulties* 20 to watch (television) 21 (tr) to sight (a hunted animal) before or during the chase [C15: from Old French *veue*, from *veoir* to see, from Latin *vidēre*] ▸ **'viewable** adj ■ **Viewdata** ('vjuː,deɪtə) n trademark an interactive form of Videotex that sends information from a distant computer along telephone lines, enabling shopping, booking theatre and airline tickets, and banking transactions to be conducted from the home ■ **viewer** ('vjuːə) n 1 a person who views something, esp television 2 any optical device by means of which something is viewed, esp one used for viewing photographic transparencies 3 law a person appointed by a court to inspect and report upon property, etc ▸ **'viewership** n ■ **viewfinder** ('vjuː,faɪndə) n a device on a camera, consisting of a lens system and sometimes a ground-glass screen, enabling the user to see what will be included in his or her photograph. Sometimes shortened to: **finder** ■ **view halloo** interj 1 a huntsman's cry uttered when the quarry is seen breaking cover or shortly afterwards ▸ n 2 a shout indicating an abrupt appearance ■ **viewing** ('vjuːɪŋ) n 1 the act of watching television 2 television programmes collectively: *late-night viewing* ■ **viewless** ('vjuːlɪs) adj 1 (of windows, etc) not

V

EXHIBIT A-13

Deep Convolutional Neural Network Features and the Original Image

Connor J. Parde[1] and Carlos Castillo[2] and Matthew Q. Hill[1] and Y. Ivette Colon[1] and
Swami Sankaranarayanan[2] and Jun-Cheng Chen[2] and Alice J. O'Toole[1]

[1] School of Behavioral and Brain Sciences, The University of Texas at Dallas, USA

[2] Department of Electrical Engineering, University of Maryland, College Park, USA

Abstract— **Face recognition algorithms based on deep convolutional neural networks (DCNNs) have made progress on the task of recognizing faces in unconstrained viewing conditions. These networks operate with compact feature-based face representations derived from learning a very large number of face images. While the learned feature sets produced by DCNNs can be highly robust to changes in viewpoint, illumination, and appearance, little is known about the nature of the face code that emerges at the top level of such networks. We analyzed the DCNN features produced by two recent face recognition algorithms. In the first set of experiments, we used the top-level features from the DCNNs as input into linear classifiers aimed at predicting metadata about the images. The results showed that the DCNN features contained surprisingly accurate information about the yaw and pitch of a face, and about whether the input face came from a still image or a video frame. In the second set of experiments, we measured the extent to which individual DCNN features operated in a view-dependent or view-invariant manner for different identities. We found that view-dependent coding was a characteristic of the *identities* rather than the DCNN features--with some identities coded consistently in a view-dependent way and others in a view-independent way. In our third analysis, we visualized the DCNN feature space for 24,000+ images of 500 identities. Images in the center of the space were uniformly of low quality (e.g., extreme views, face occlusion, poor contrast, low resolution). Image quality increased monotonically as a function of distance from the origin. This result suggests that image quality information is available in the DCNN features, such that consistently average feature values reflect coding failures that reliably indicate poor or unusable images. Combined, the results offer insight into the coding mechanisms that support robust representation of faces in DCNNs.**

I. INTRODUCTION

Face recognition algorithms based on convolutional neural networks and deep learning show considerable robustness to changes in imaging parameters (e.g., pose, illumination, and resolution) and facial appearance (e.g., expression, eyewear). This robustness accounts for the impressive gains made by CNNs on the problem of unconstrained face recogniton [1], [2], [3], [4], [5], [6]. Performance on datasets such as LFW [7], [8], IJB-A [9], [10], and Mega-Face [11] offer evidence that face recognition by machines can, in some cases, approach human performance [1]. Indeed, human recognition of familiar faces (e.g., friends, family) operates in highly unconstrained environments and over changes in appearance and age that can span decades. This kind of performance remains a goal of automated face recognition systems.

Although humans remain a proof-of-principle that highly invariant face recognition is possible, the underlying nature of the face representation that supports invariance in humans is poorly understood. The nature of the representation captured in DCNN features is similarly elusive. The goal of this paper is to characterize the features that emerge in a DCNN trained for face recognition so as to better understand why they are robust to yaw, pitch, and media type (still image or video frame). The approach we take is to first examine the extent to which the "robust" feature sets that emerge in a CNN *retain* information about the original images. As we will see, DCNNs that show considerable robustness to pose and media type retain detailed information about the images they encode at the deepest and most compact level of the network. Second, we explore the view-dependency and media-dependency characteristics of DCNN features. Third, we examine cues pertaining to image quality within the structure of the feature space.

II. BACKGROUND AND PROBLEM

The problem of image-invariant face perception has been studied for decades in both computer vision [12] and psychology. Traditionally, two classes of models have been considered: a.) representations that capture 3D facial structure and b.) representations based on collections of 2D, image-based views of faces. The former can enable specification of appearance across arbitrary affine and non-affine transformations. The latter can show invariance in any given instance via interpolation to image representations taken in conditions similar to that of the probe image. Notably, this requires "experience" with enough diverse views to be successful across a range of possible probes. Active appearance models [13] comprise an intermediary class, which relies on class-based knowledge of faces, including 3D structure and reflectance-map information for many examples. Although these models can achieve impressive performance in computer graphics representations made from single images, they are not practical for face recognition as they are computationally intense and require high quality, 3D data on diverse classes of faces.

The recent gains made in face recognition can be tied both to the computational power of DCNNs and to the quality and quantity of the training data now available from web-scraping. In theory, the goal of a DCNN is to develop an invariant representation of an individual's face through exposure to a wide variety of images showing that person in different settings, with different poses, and in images that vary in quality. Given enough data, it is expected that the

network will be able to learn a representation of an individual that does not rely on these non-static, image-level attributes. Instead, the intent is that the learned features represent the invariant information in a face that makes the face unique.

The fact that DCNNs support robust recognition across image transformation does not preclude the possibility that the features used to code faces in these networks also retain information about the image properties. Rather, DCNNs may succeed across appearance-related and image-related variation by incorporating both identity and image parameters into the face codes. This code may support the separation of image and identity for identity verification. This separation may ultimately be achieved at a post-DCNN stage via another type of classifier that operates on image or person representations extracted from the deepest, most compact layer of the DCNN.

The motivation for the present work came from visualizing the way *single identities* cluster in a low-dimensional space derived from the top-level features produced by two recent DCNNs [9], [10]. These DCNNs were developed to work on the Janus CS2 dataset, an expanded version of the IJB-A dataset [16]. We describe the architecture of the two DCNNs in detail in the methods section. For present purposes, this visualization was done by applying t-Distributed Stochastic Neighbor Embedding (t-SNE) [14] to the top level features of each network. t-SNE is a dimensionality reduction technique that uses stochastic probability methods to preserve the high-dimensional Euclidean distances between data points while embedding them in a low-dimensional space. We visualized single identities that had large numbers of images available in the Janus CS2 dataset. Figure 1 shows the t-SNE space for the top level features of 140 CS2 images of Vladimir Putin, extracted from the two DCNNs. Both plots exhibit roughly separable clusters of profile and frontal images of the subject. The blue curves were hand-drawn onto the visualizations to indicate the position of an approximate border.

The clustering suggests that the top-level features produced by both of these DCNN networks preserve salient, view-related information captured in the original image, while still clustering by identity. More generally, this suggests that DCNNs contain a deeper-than-expected representation of the original image in their top-level features. Notably, the clustered images of Putin still varied substantially in other appearance- and image-based attributes (e.g., age, illumination).

In what follows, we quantify the clustering behavior of image-based attributes in these two DCNNs. This paper is organized as follows. In Section III, we present the networks and the datasets analyzed. In Section IV we use the top-level features of the DCNNs as input into linear classifiers aimed at predicting metadata about the images including yaw, pitch, and media type (still image or video). In Section V, we analyzed the extent to which top-level features operate invariantly across viewpoint and media type. In Section VI, we examine the top-level feature space of image representations in the context of image quality.
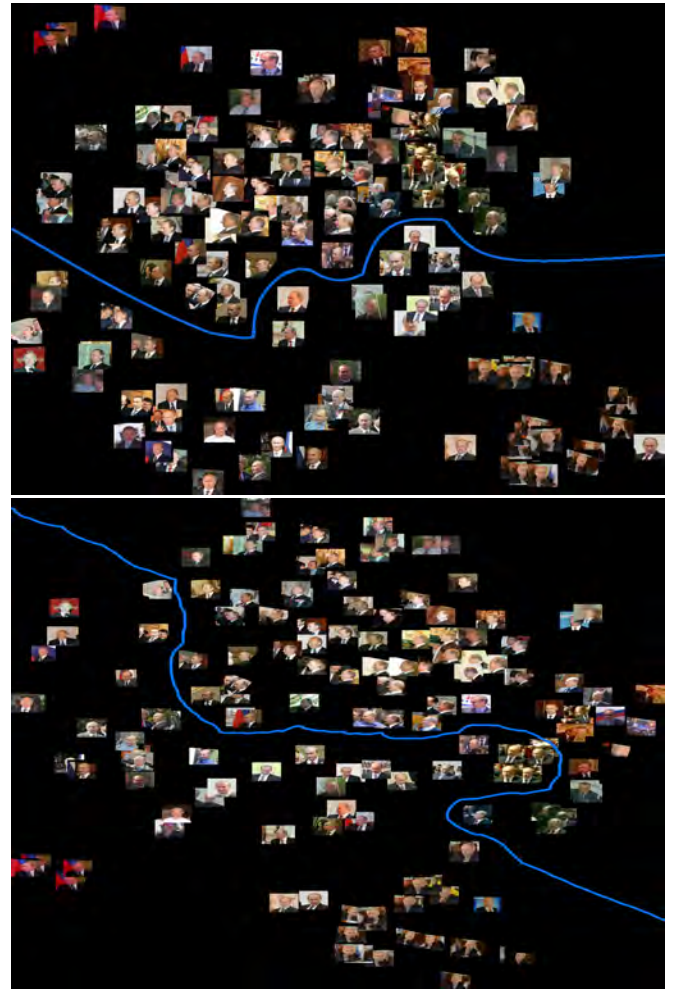

Fig. 1. These figures show the t-SNE visualization of the top level DCNN features for 140 images of Vladimir Putin from the Janus CS2 dataset. The visualizations are based on the 320 top-level DCNN features from Network A [9] (top) and the 512 top-level DCNN features from Network B [10].

III. General Methods

A. Description of DCNN's

We analyzed the feature-space produced by two DCNNs (Network A, [9]; Network B, [10]) using the Janus CS2 dataset. Both networks approach the problem by constructing a feature-based representation of all input images using a DCNN. Full details on these networks are available elsewhere. For present purposes, we discuss only the training of these two networks since our analysis focuses only on the top-level features produced by either network.

The base architectures of the DCNNs appear in Tables I and II. In both networks, parametric ReLU (PReLU) were used as the activation function. In Network A, a learned feature space was developed from scratch and produced a 320-dimensional feature vector for each input image. The second network (Network B) builds upon the AlexNet model [15] and assigns each input image a 512-dimensional feature vector. At its lower levels, Network B initially assigns weights based on the values generated by AlexNet and then trains its higher layers using the CASIA-Webface database.

TABLE I

NETWORK A

Name	Filter Size/Stride	Output	Parameters
conv11	3x3x1/1	100x100x32	.28K
conv12	3x3x32/1	100x100x64	18K
pool1	2x2/2	50x50x64	
conv21	3x3x64/1	50x50x64	36K
conv22	3x3x64/1	50x50x128	72K
pool2	2x2/2	25x25x128	
conv31	3x3x128/1	25x25x96	108K
conv32	3x3x96/1	25x25x192	162K
pool3	2x2/2	13x13x192	
conv41	3x3x192/1	13x13x128	216K
conv42	3x3x128/1	13x13x256	288K
pool4	2x2/2	7x7x256	
conv51	3x3x256/1	7x7x160	360K
conv52	3x3x160/1	7x7x320	450K
pool5	7x7/1	1x1x320	
dropout (40%)		1x1x320	
fc6		10548	3296K
softmax cost		10548	
total			5006K

TABLE II

NETWORK B

Layer	Kernel Size/Stride	Parameters
conv1	11 x 11/4	35K
pool1	3 x 3/2	
conv2	5 x 5/2	614K
pool1	3 x 3/2	
conv3	3 x 3/2	885K
conv4	3 x 3/2	1.3M
conv5	3 x 3/1	2.3M
conv6	3 x 3/1	2.3M
conv7	3 x 3/1	2.3M
pool7	6 x 6/2	
fc6	1024	18.8M
fc7	512	524K
fc8	10548	10.8M
Softmax Loss		Total 39.8M

Network A also uses CASIA-Webface for training and does so for all layers in the network.

B. CS2 Dataset

The images were sourced from the CS2 dataset. The dataset includes 25,800 images of 500 subjects. CS2 is an expanded version of the IARPA Janus Benchmark A (IJB-A) [16], a publicly available "media in the wild" dataset. Some key features of the IJB-A dataset include: full pose variation, a mix of images and videos, and a wider demographic variation of subjects than is available in the LFW dataset. The dataset was developed using 1,501,267 crowd sourced annotations. Baseline accuracies for both face detection and face recognition from commercial and open source algorithms are available in [16].

The original IJB-A included metadata from crowd-sourcing. Here we used metadata provided by the Hyperface system described in [17]. The Hyperface system provides key-point locations to aid in face detection, as well as estimated measurements of face pose (yaw, pitch, and roll).

Of the 25,800 items in the CS2 dataset, we omitted 1,298

TABLE III

YAW AND PITCH PREDICATION ACCURACY

Network	Yaw	Pitch
A	+/-8.06 degs. (*sd.* 0.078)	77.0% correct
B	+/-8.59 degs. (*sd.* 0.071)	71.5% correct

items from our analysis. This was due to either Network A's or Network B's inability to compute features for one of these images, or Hyperface's inability to compute the pose of the subject within an image. This left us with 24,502 items that could be considered when training classifiers to predict each metadata attribute of interest.

IV. PREDICTING IMAGE-RELATED METADATA FROM THE DCNN FEATURES

Each experiment described in this section consisted of a bootstrap test of metadata prediction based on the top-level feature encodings from Network A and B. Predictions were computed using a linear discriminant analysis (LDA) classifier, with 20 iterations of the bootstrap test for each metadata attribute. For each iteration, we randomly selected 18,000 items to use as training data. We tested the classifier on the remaining 6,502 items. The results reported on prediction accuracy are averaged across the 20 bootstrapped iterations.

A. Predicting Yaw

The yaw values provided by the Hyperface system for the CS2 dataset describe the yaw angle of the face in an image, measured in degrees, and varying from -90 (left profile) to +90 (right profile), with 0 indicating a frontal pose. For both networks, pre-processing steps were performed to mirror all left-facing images, thereby limiting the yaw range to only positive values. Therefore, we used the absolute value of the yaw scores provided by Hyperface as output for the classifier. In each bootstrap iteration, a classifier was trained to predict the Hyperface yaw values from the DCNN features. Prediction accuracies for both Networks A and B appear in Table III and are surprisingly high, to within less than 9 degrees, and are consistent across bootstrap iterations.

B. Predicting Pitch

Pitch estimates for the dataset were provided by Hyperface and measured in degrees. A positive pitch score indicates an upward looking face, a negative pitch indicates a downward looking face, and a score of 0 indicates a face looking directly at the camera. Given that the majority of images in the CS2 dataset showed faces with a relatively centered pitch, pitch was coded categorically for this experiment as *centered* and *deviating* pitch. Centered pitch was defined as all values between -8 and +8 degrees. Deviating pitch was defined as all values outside of the centered range.

Using the DCNN features as input, we predicted whether each image in the CS2 data set showed a face with centered pitch or deviating pitch. Predictions on the test data were continuous values from 0 (centered) to 1 (deviating). These were rounded to the nearest integer (0 or 1) to obtain the

TABLE IV

MEDIA TYPE

Network	Media Type
A	87.1% (sd. 0.004)
B	93.3 % (sd. 0.002)

prediction values. The results appear in Table III, reported as percent correct. As with yaw, pitch category prediction from the DCNN features was unexpectedly accurate (77.0% and 71.5% correct) for Networks A and B, respectively.

C. Predicting Media Type

The media type of each image was provided in the CS2 dataset. Each image originated as either a still photograph or a video frame. An image's media type might be considered a proxy-measure for some aspects of image quality. In the CS2 dataset, the images that originated as still photographs were typically better illuminated and had higher resolution. The images that originated as video frames tended to come from lower-quality data sources such as CCTV footage.

We assigned a score of 1 to all images in the CS2 dataset that originated as still photographs, and a score of 0 to all images that originated as video frames. We then applied the bootstrapped classification method to predict media type from the CNN features. The predictions for our test data were continuous values from 0 to 1. These were rounded to the nearest integer (0 or 1) to obtain the prediction values. The results appear in Table IV, reported as percent correct. Predictions using the DCNN features were highly accurate and consistent for both networks.

D. Interim summary

The classification experiments showed that metadata from individual images, including yaw, pitch, and media type, was available in the top level DCNN features of both networks.

In the next section the goal was to analyze the extent to which individual features operate invariantly, or at least robustly, across pose and media type.

V. CNN FEATURES AND INVARIANCE: ARE FEATURES INVARIANT OR ARE PEOPLE INVARIANT?

A. View (In)variance Coding

Beginning with view, we developed an index of feature robustness across frontal and profile poses. We approached the problem as follows. First, we sub-selected identities in the database ($n = 38$) for which there were at least 20 frontal images and 20 profile images. Second, within each of these identities, for each of the 320 DCNN features in Network A, we computed a t-test to determine whether the feature's values from frontal images of that individual differed significantly from the feature's values from profile images. We set the alpha level for statistical significance at 0.000156[1]. The resultant p-values act as an of index

[1]This is a two-tailed alpha level of 0.05, Bonferroni corrected for 320 multiple comparisons.
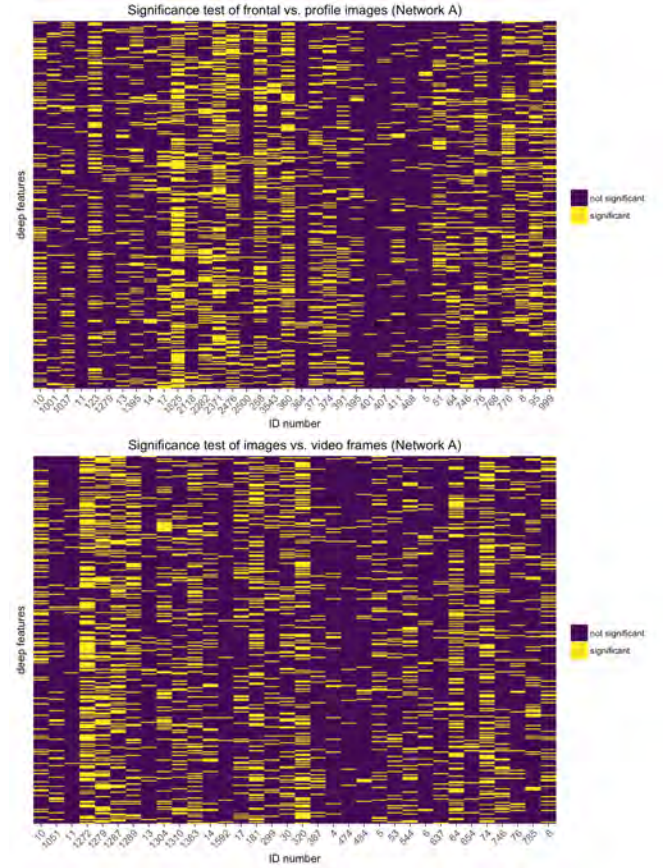

Fig. 2. Heat map illustration of view-dependent DCNN features for Network A displayed for each identity in the database with at least 20 frontal and 20 profile images (top). Heat map illustration of the quality-dependent top-level DCNN features for each identity in the database with at least 20 still images and 20 video frames (bottom).

of feature invariance for an individual. The results of this analysis are displayed in the top panel heat map in Figure 2 and are surprising. In the figure, individual identities are displayed across columns and individual features are displayed across rows. We anticipated that individual features would consistently code identities in either a view-dependent or view-invariant way. This would have produced horizontal bands in the heat map, suggesting the consistency of a feature across identities. Instead we found the inverse, with individual identities being coded in either a view-dependent or view-invariant way across features. This is indicated by the vertical lines of significant features in the heat map. More formally, the percentage of features that differentiated faces by viewpoint for an individual was as high as 55.31%. Individual features did not consistently code in a view-dependent or view-independent manner.

To interpret these heat maps, we visualized the most- and least-differentiated identities by selecting the most strongly banded columns from the heat map. Two examples of non-differentiated identities appear in Figure 3 and show Bono and Pres. George W. Bush. For Bono, 90.31% of the 320 features were undifferentiated as a function of view; for Bush, 97.5 % were undifferentiated. These clusters show

Fig. 3.   Image clusters of two individuals (Bono and G. Bush) who were both coded with a majority of view-independent features (312 and 289 of 320 respectively). These clusters show mixed viewpoints aligned closely, which may correspond to distinctive features (e.g. Bono's sunglasses) that are easy to detect across variable views.

mixed viewpoints aligned closely–possibly reflecting the presence of distinctive identity features that are easy to detect in any view (e.g. Bono's oddly tinted sunglasses). In visualizing identities with the most differentiated features, however, many subjects show strongly separated clusters, each of which shows a small range of similar views. This latter pattern resembles what we saw in Figure 1 for Vladimir Putin. The main point, though, is that it is the *identity* that determines whether the features will operate in a view-dependent or view-invariant manner. Some identities are marked most strongly by characteristics which are static across shifts in pose, while others are marked by the way certain traits appear when seen from different viewing angles.

To determine the extent to which the nature of an identity code (view-invariant or non-invariant) affects performance in a face recognition algorithm, we conducted the following experiment. We selected the 7 identities coded most invariantly over view-change. Next we compared the performance of Network A on template comparisons comprised of pairs of these 7 identities against templates comprised of all other identity pairs. Note that a template is defined as a variably sized set of images and video frames of an individual identity. The contents of the templates were specified by the Janus protocol. The results appear in Figure 5 and show a strong advantage for recognizing identities that can be coded invariantly, over those in which feature values dissociate for frontal and profile images.

B. Media Type (In)variance Coding

We repeated the same approach from the previous section to examine the way media type is coded across features and individuals, developing an index of feature robustness across still images and video frames. First, we sub-selected identities in the database ($n = 34$) for which there were at least 20 still images and 20 video frames. Second, within each of these 34 identities, for each of the 320 top-level DCNN features produced by Network A, we computed a t-test to determine whether a feature's value for still images of that individual differed significantly from that feature's value for video frames. We again set the alpha level for statistical significance at 0.000156. In this case, the p-values act as an index of the feature's invariance for coding an
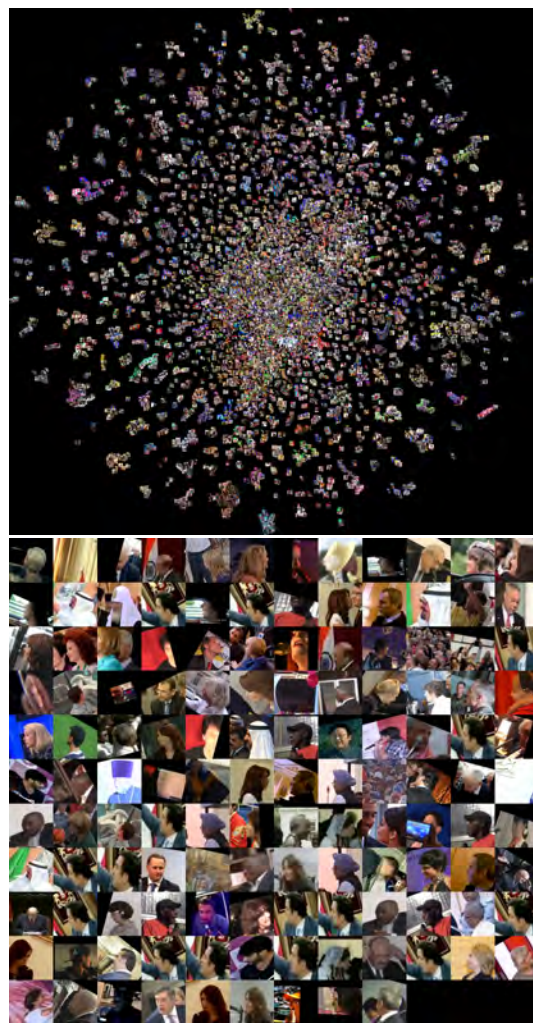

Fig. 4.   Results of t-SNE applied to the DCNN top level features of Network A for all 24,502 images (top). An array of the 129 images closest to the center of the space (0.05%) in Network A. The upper-left image is the image closest to the center, and each image's distance from the center grows as you progress across the rows (bottom).

individual in a still photograph versus in a video frame. The results of this analysis are displayed in the heat map in Figure 2 (bottom panel) and echo what is seen in the heat map distinguishing frontal and profile views. Individual identities tend to be coded in either a media-dependent or media-independent manner.

VI. WHEN DCNN FEATURES FAIL THEY LEAVE A TRAIL

We returned to the use of t-SNE to visualize the feature spaces of our two recognition networks. This time, rather than analyzing the feature space for a single individual, we applied t-SNE to the DCNN top level features for all 24,502 images (see Figure 4, top). This was used as an exploratory analysis to help us visualize the DCNN feature space in more detail. The primary insight gained from this visualization is that the images located near the center appear to be of extremely poor "quality", where quality refers to a wide range of issues that would make the person in the image difficult to detect or identify. We therefore examined the images in order
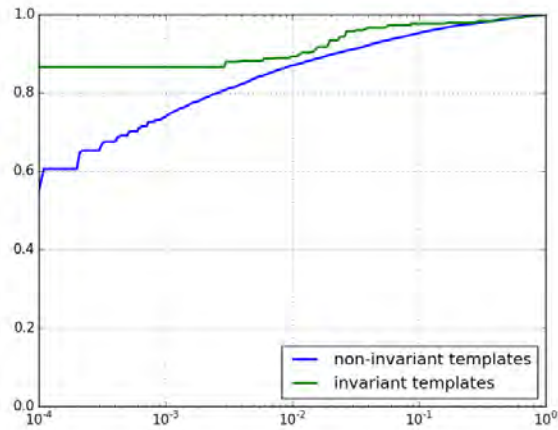
Fig. 5.   Identity verification performance of Network A for template pairs where both identities are coded view-invariantly versus for all other template pairs. View-invariance of an identity is characterized by feature values across its images that do not dissociate for frontal and profile views.

of closeness to the center of the raw feature space, using the origin of the feature space as the center. Figure 4 (bottom) shows an array of the 129 images closest to the center of the space (0.05%) in Network A, arranged across the rows and starting from the image closest to the center. As seen in the array, the images closest to the center of the feature space are affected by a range of problems including extreme views, strong occlusion, blurring, distortion, and lack of an identifiable face.

Does distance from the center of the DCNN feature space index image quality? To examine this, we pulled images from different distances to the center of the space. We ranked the images according to their distance from the origin. Figure 6 shows 258 sampled images at the 20th, 50th, and 90th percentiles of these ranked distances. This figure illustrates that face quality seems to increase with distance from the center of the DCNN feature space.

VII. CONCLUSIONS

The three analyses we carried out yielded the following results. First, DCNN top-level features retain a surprising amount of information about the original input imagery. Yaw, pitch, and media type were readily available in the top-level DCNN codes, and could be classified with high accuracy.

Second, in characterizing the extent to which individual features coded view-dependent or view-invariant information about faces, we found that view-dependent coding was a characteristic of the *identities* rather than the features. This might imply that some identities in this dataset present with appearance characteristics that are easy to detect and code across views, whereas other identities have few appearance characteristics that can generalize across view. This data-dependent code is intriguing in that it suggests that DCNNs and the human visual system alike might need to exploit both types of codes to operate efficiently and accurately in unconstrained viewing conditions. The same general finding
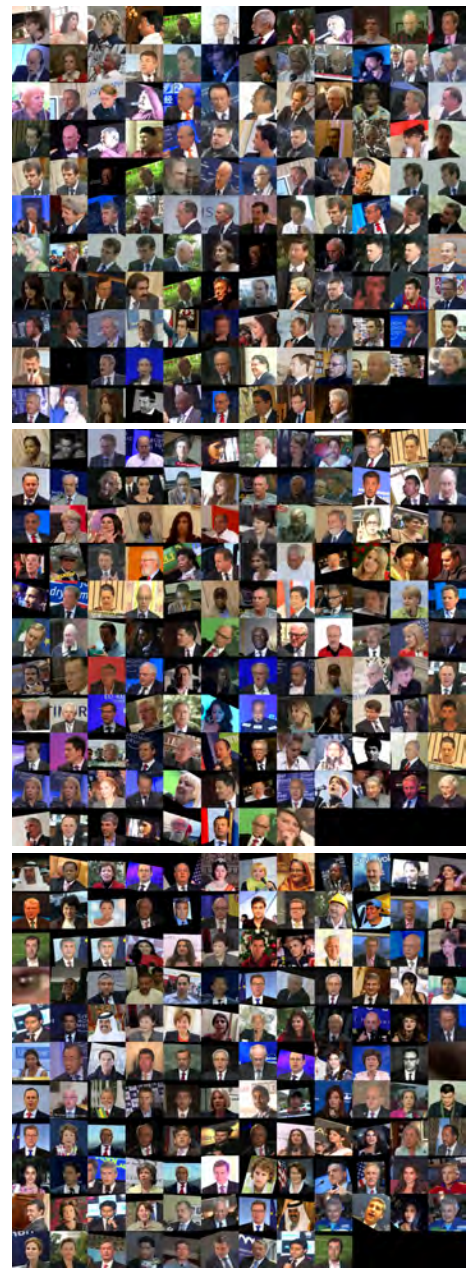

Fig. 6.   Images (n=129) sampled at the 20th (top), 50th (middle), and 90th (bottom) percentiles of ranked distances from the origin. Face image quality seems to increase with distance from the center of the DCNN feature space.

of data-dependency held for media type as well, with some identities having consistent codes across media types and others having disparate codes.

Finally, we found an unexpected index of image quality in the DCNN space. This took the form of distance from the origin of the space. The clustering of poor quality images was notable in that the low quality emanated from many distinct sources. This allowed for a more generic definition of images with limited or unusable information about identity. Because low quality images cluster around the origin and quality increases with distance from the origin, we might speculate that strong DCNN feature scores reflect

robust identity information. This suggests a new method for screening out poor quality imagery in DCNNs.

In summary, a more in-depth look at DCNN features gave insight into the nature of the image information in these top-level compact feature codes. These analyses point to data-dependent flexibility in the type of codes that emerge at the top level features, as well as the possibility of isolating bad data from better quality imagery.

VIII. ACKNOWLEDGMENTS

REFERENCES

[1] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "Deepface: Closing the gap to human-level performance in face verification," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2014, pp. 1701–1708.

[2] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep face recognition," *Proceedings of the British Machine Vision*, 2015.

[3] K. Simonyan, O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Fisher vector faces in the wild," in *BMVC*, vol. 60, no. 2, 2013.

[4] F. Schroff, D. Kalenichenko, and J. Philbin, "Facenet: A unified embedding for face recognition and clustering," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.

[5] G. Hu, Y. Yang, D. Yi, J. Kittler, W. J. Christmas, S. Z. Li, and T. M. Hospedales, "When face recognition meets with deep learning: an evaluation of convolutional neural networks for face recognition," *CoRR*, vol. abs/1504.02351, 2015. [Online]. Available: http://arxiv.org/abs/1504.02351

[6] G. B. Huang, H. Lee, and E. Learned-Miller, "Learning hierarchical representations for face verification with convolutional deep belief networks," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. IEEE, 2012, pp. 2518–2525.

[7] G. B. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled Faces in the Wild: a database for studying face recognition in unconstrained environments," University of Massachusetts, Amherst, Tech. Rep. 07-49, 2007.

[8] N. Kumar, A. C. Berg, P. N. Belhumeur, and S. K. Nayar, "Attribute and Simile Classifiers for Face Verification," in *Proceedings of the 12th IEEE International Conference on Computer Vision (ICCV)*, October 2009.

[9] J.-C. Chen, "Unconstrained face verification using deep cnn features," *arXiv preprint arXiv:1508.01722*, 2016.

[10] S. Sankaranarayanan, "Triplet probabilistic embedding for face verification and clustering," *arXiv preprint arXiv:1604.05417*, 2016.

[11] D. Miller, I. Kemelmacher-Shlizerman, and S. M. Seitz, "Megaface: A million faces for recognition at scale," *arXiv preprint arXiv:1505.02108*, 2015.

[12] K.-K. Sung and T. Poggio, "Example-based learning for view-based human face detection," *IEEE Trans. PAMI*, vol. 20, pp. 39–51, 1998.

[13] V. Blanz and T. Vetter, "A morphable model for the synthesis of 3d faces," in *Proceedings, SIGGRAPH'99*, 1999, pp. 187–194.

[14] L. van der Maaten, "Visualizing data using t-sne," *Journal of Machine Learning Research*, vol. 9, 2008.

[15] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[16] B. F. Klare, B. Klein, E. Taborsky, A. Blanton, J. Cheney, K. Allen, P. Grother, A. Mah, M. Burge, and A. K. Jain, "Pushing the frontiers of unconstrained face detection and recognition: Iarpa janus benchmark a," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015, pp. 1931–1939.

[17] R. Ranjan, V. M. Patel, and R. Chellappa, "Hyperface: A deep multi-task learning framework for face detection, landmark localization, pose estimation, and gender recognition," *arXiv preprint arXiv:1603.01249*, 2016.